

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 08:14:28, Search time 347 Seconds
(without alignments)
4500.606 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186

Sequence: 1 caggtgaltatcctaataa.....caccacaacacttaag 3186

Scoring table: IDENTITY NUC

Searched: 4161431 seqs, 245089505 residues

Total number of hits satisfying chosen parameters: 8322862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.New.*
1: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq2:*
8: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq3:*
9: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
10: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*

*Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97.2	3.1	1290	6	US-10-750-185-43433
2	59.4	1.9	10968	7	US-11-075-185-35
3	59.4	1.9	78869	7	US-11-075-185-1
4	55.2	1.7	1280	6	US-10-858-730-47
5	54.8	1.7	116856	7	US-11-143-980-1
6	54.2	1.7	5679	7	US-11-075-185-36
7	54.2	1.7	14172	7	US-11-075-185-2
8	53.2	1.7	5679	7	US-11-075-185-36
9	53.2	1.7	78869	7	US-11-075-185-1
10	52.2	1.6	2302	7	US-11-110-082-20
11	52	1.6	2079	7	US-11-045-802-25
12	52	1.6	7006	6	US-10-821-234-218
13	52	1.6	18138	6	US-10-995-561-13385
14	51.6	1.6	996	6	US-10-432-483-22
15	51.6	1.6	6941	6	US-10-432-483-49
16	51.4	1.6	2148	6	US-10-821-234-29
17	51.2	1.6	3513	6	US-10-858-730-142
18	50.8	1.6	3381	7	US-11-075-185-33
19	50.8	1.6	170995	7	US-11-121-086-35
20	50.6	1.6	191684	7	US-11-121-086-2
21	50.4	1.6	171936	6	US-10-933-025-24
22	50	1.6	2033	6	US-10-995-561-55
23	50	1.6	3305	7	US-11-017-550-68

24	50	1.6	12391	7	US-11-017-550-67	Sequence 67, Appl
25	50	1.6	12591	6	US-10-995-561-13415	Sequence 13415, A
26	50	1.6	14023	6	US-10-995-561-13221	Sequence 13221, A
27	50	1.6	199130	6	US-10-995-561-13233	Sequence 13233, A
28	49.8	1.6	2584	6	US-10-821-234-836	Sequence 836, App
29	49.6	1.6	1593	6	US-10-858-730-138	Sequence 138, App
30	49.4	1.6	3408	6	US-10-858-730-40	Sequence 40, Appl
31	49.4	1.6	11070	7	US-11-075-185-34	Sequence 34, Appl
32	48.4	1.6	96128	6	US-10-995-561-13197	Sequence 13197, A
33	48.4	1.5	61718	6	US-10-995-561-13326	Sequence 13326, A
34	48.6	1.5	11119	7	US-11-143-980-11	Sequence 11, Appl
35	48.6	1.5	116856	7	US-11-143-980-1	Sequence 1, Appl
36	48.4	1.5	4509	7	US-11-087-100-5	Sequence 5, Appl
37	48.4	1.5	4509	7	US-11-087-084-5	Sequence 5, Appl
38	48.4	1.5	4509	7	US-11-087-085-5	Sequence 5, Appl
39	48.4	1.5	4982	6	US-10-276-233A-17	Sequence 17, Appl
40	48.4	1.5	11070	7	US-11-075-185-34	Sequence 34, Appl
41	48.2	1.5	1719	6	US-10-848-724-4	Sequence 4, Appl
42	48	1.5	10968	7	US-11-075-185-35	Sequence 35, Appl
43	47.8	1.5	888	6	US-10-858-730-183	Sequence 183, App
44	47.8	1.5	944	6	US-10-667-295-150	Sequence 150, App
45	47.8	1.5	8651	6	US-10-432-483-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT 1
US-10-750-185-43433
; Sequence 43433, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR IMPROVING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 43433
; LENGTH: 1290
; TYPE: DNA
; ORGANISM: Bovine 19866881067936
US-10-750-185-43433

Query Match          3.1% Score 97.2; DB 6; Length 1290;
Best Local Similarity 75.9%; Pred. No. 6.1e-12;
Matches 120; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 993 AGCGCATGTTCTTTTAAAGTTTAACTTTCTGTCGATCCACGCTATTAC 1052
DB 478 AACGCGATTTCTTTTGAAGTTCAATTAACGACGATCCACCCGCTATC 537
QY 1053 AATATTTTGTGATGATGATTTTGGCGATCCCATCATCTGAGATTCAAGAGGCAAA 1112
DB 538 AACGTTCTGTGAAGTGTGCTGCGGACCCCAACACTGCGCTTCAGGGGAGCAAA 597
QY 1113 TGGGTTCTTGGCGCAAGCGGACCAATGTCAGG 1150
DB 598 TGGGTCACCTGCGGAAAGCGGACATTAATGTCAGG 635

RESULT 2
US-11-075-185-35
; Sequence 35, Application US/11075185
```

Publication No. US20050266434A1
GENERAL INFORMATION:
APPLICANT: REEVES, CHRISTOPHER D
APPLICANT: JULIEN, BRIAN
APPLICANT: REID, RALPH
TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
FILE REFERENCE: 010099.03
CURRENT APPLICATION NUMBER: US/11/075,185
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/551,103
PRIOR FILING DATE: 2004-03-08
PRIOR APPLICATION NUMBER: US 60/568,290
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.3
SEQ ID NO 35
LENGTH: 10968
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-11-075-185-35

Query Match 1.9%; Score 59.4; DB 7; Length 10968;
Best Local Similarity 43.6%; Pred. No. 0.0023;
Matches# 325; Conservative 0; Mismatches 411; Indels 9; Gaps 1;

QY 1708 CGCCGACAGAGCCGAGACCGCGGCGCCCTCGCCGCAACGCTGTTGTGACGCCG 1767
DB 6728 CGCACAACCGCGGAGGCGCGCGCGCGCTGCGCTGACGCTGCCGCTCTCTGCG 6787
QY 1768 CCACACACCGCTGACCTTCGCGGCTCGCCCTATGACACGCGACGCACTTCGCGGCA 1827
DB 6788 GTCGCGAGAGCGCGGCGCTGCGCGCCGACGCGCGCGCGCTGCGACGACCTGCGCGCC 6847
QY 1828 ACCGCGCACGCTGCTCTTACGCGCGCGCGCGGTGAAGCGCTGCCGCTGCAAGCTG 1887
DB 6848 ACCCGGACGAGCGGCTGCTGACGCTGCGCGCGCGCTGCGCACGACGCGCACTGCG 6907
QY 1888 CAGGCTGACCTGCGCGCGCGCGCTGCGCTACTAGCGGACCGCGCGCGCGCGCGCA 1947
DB 6908 CCACGCGGCTCGCTGCG 6967
QY 1948 GTCCCGCAGTACTGCGGACCAAGTCGCGGCTGCGGCTGCGCTGCGCTGCGCGCA 2007
DB 6968 TTGCGGAGTTGCGCTCAAGC-----GCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7018
QY 2008 CCG 2067
DB 7018 GTCCCGGCGAGCG 7078
QY 2068 TGGCGCGCGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2127
DB 7078 CCGGCACTGCG 7138
QY 2128 TGTCCGATTCAGCTGATGAGAGCGCTCTCTGATCAAGTCAATGCACTTCACGCACT 2187
DB 7138 CTTGCGCGAGCTGACCG 7198
QY 2188 CCGGCGATTTACAGAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2247
DB 7198 CCGGCTGCGAGCG 7258
QY 2248 AGAGTTGCTCCCGCTCAAGAGCGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 2307
DB 7258 CTCTGAGGTGCGCGCTCTACCGCGCGAGTGGAGTCTGCGCGCGCGCGCGCGCGCG 7318
QY 2308 CCAGGACATTAGCGGCTACTAGCTTCTACGACAGTAAAGCGCGCGCGCGCGCGCG 2367
DB 7318 TCGGCGACAGATCGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7378
QY 2368 CGGCG 2427
DB 7378 ACGCTTCCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7438

QY 2428 CCCACACTCTCTTGTGACGCCCA 2452
DB 7439 CCATGCGCTTCTGAGGCCACCA 7463

RESULT 3
US-11-075-185-1
Sequence 1, Application US/11075185
Publication No. US20050266434A1
GENERAL INFORMATION:
APPLICANT: REEVES, CHRISTOPHER D
APPLICANT: JULIEN, BRIAN
APPLICANT: REID, RALPH
TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
FILE REFERENCE: 010099.03
CURRENT APPLICATION NUMBER: US/11/075,185
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/551,103
PRIOR FILING DATE: 2004-03-08
PRIOR APPLICATION NUMBER: US 60/568,290
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1
LENGTH: 78869
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-11-075-185-1

Query Match 1.9%; Score 59.4; DB 7; Length 78869;
Best Local Similarity 43.6%; Pred. No. 0.0047;
Matches 325; Conservative 0; Mismatches 411; Indels 9; Gaps 1;

QY 1708 CGCCGACAGAGCCGAGAGCCCGCGCGCCCTCGCCGCAACGCTGTTGTGACGCCG 1767
DB 2195 CGCAGAGCGCGCGGAGCGCGCGCGCGCTGCGCTGACGCTGCCGCTCTCTGCG 22054
QY 1768 CCACACACCGCTGACCTTCGCGGCTCGCCCTATGACACGCGACGCACTTCGCGGCA 1827
DB 2205 GTCGCGAGAGCGCGGCGCTGCGCGCCGACGCGCGCGCGCGCTGCGACGCTGCCGCGCC 22114
QY 1828 ACCGCGCACGCTGCTCTTACGCGCGCGCGCGGTGAAGCGCTGCCGCTGCAAGCTG 1887
DB 2215 ACCCGGACGAGCGGCTGCTGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22174
QY 1888 CAGGCTGACCTGCGCGCGCGCGCTGCGCTACTAGCGGACCGCGCGCGCGCGCGCA 1947
DB 22175 CCACGCGGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22234
QY 1948 GTCCCGCAGTACTGCGGACCAAGTCGCGGCTGCGGCTGCGCTGCGCTGCGCGCA 2007
DB 2225 TTGCGGAGTTGCGCTCAAGC-----GCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22285
QY 2008 CCG 2067
DB 22286 CCGCGGAGGAGCG 22345
QY 2068 TGGCGCGCGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2127
DB 22346 CCGGCACTGCG 22405
QY 2128 TGTCCGATTCAGCTGATGAGAGCGCTCTCTGATCAAGTCAATGCACTTCACGCACT 2187
DB 22406 CTCTGAGGTGCGCGCTCTACCGCGCGAGTGGAGTCTGCGCGCGCGCGCGCGCGCG 22465
QY 2188 CCGGCGATTTACAGAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2247
DB 22466 CCGGCTGCGAGCGCGCGCGCGCTGCTGACGACGAGCTGCGCGCGCGCGCGCGCG 22525
QY 2248 AGAGTTGCTCCCGCTCAAGAGCGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 2307
DB 22526 CTCTGAGGTGCGCGCTCTACCGCGCGAGTGGAGTCTGCGCGCGCGCGCGCGCGCG 22585

QY 2308 CCAAGACATTAAGCGGCTACTATGCTTCTACTGCAAGCTAAGCCCGCCCTGACCCGCC 2367
 DB 22586 TCGGACACAGCATCGCGGAGCTCGCGCGCCCAATGCGCGGCTGCTGACCTTGGCCG 22645
 QY 2368 CGGCCCCCGCGGCGCCCGGAGCCCGGAGCCCGCTTCAAGCTTCTTCCCACTCGGCT 2427
 DB 22646 AGCGCTCGGCTTGGTGTGCGCGCGCGCGGCGGCTCATGAGGCGCTTCCCGCTTGGCGGCG 22705
 QY 2428 CCCCACTCTCTCTTTCGCGCACCA 2452
 DB 22706 CCATGGCGCTCGTTCAGGCGCACCA 22730
 RESULT 4
 us-10-858-730-47
 ; Sequence 47, Application US/10858730
 ; Publication No. US2005025568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bailey, Richard B.
 ; APPLICANT: Blomquist, Paul
 ; APPLICANT: Dotten, Reed
 ; APPLICANT: Driggers, Edward M.
 ; APPLICANT: Madden, Kevin T.
 ; APPLICANT: O'Leary, Teasica
 ; APPLICANT: O'Toole, George
 ; APPLICANT: Trueheart, Joshua
 ; APPLICANT: Walbridge, Michael J.
 ; APPLICANT: Yorgey, Peter S.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
 ; FILE REFERENCE: 14184-030001
 ; CURRENT APPLICATION NUMBER: US/10/858,730
 ; PRIOR FILING DATE: 2004-06-01
 ; PRIOR APPLICATION NUMBER: US 60/475,000
 ; PRIOR FILING DATE: 2003-05-30
 ; PRIOR APPLICATION NUMBER: US 60/551,860
 ; NUMBER OF SEQ ID NOS: 364
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 1290
 ; TYPE: DNA
 ; ORGANISM: Streptomyces coelicolor
 ; US-10-858-730-47
 Query Match 1.7%; Score 55.2; DB 6; Length 1290;
 Best Local Similarity 45.3%; Pred. No. 0.0084;
 Matches 327; Conservative 0; Mismatches 383; Indels 12; Gaps 3;
 QY 1502 CGGCTGTGACATGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 1561
 DB 51 CGGCTCAAGGTGCGGCGGATGACGACGACGCGCGGCTGACCGGCTGACCGGCTGACCGGCT 110
 QY 1562 GCGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 1621
 DB 111 GCGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 170
 QY 1622 GCGCAAGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 1681
 DB 171 GCGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 230
 QY 1682 CGGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 1741
 DB 231 CGGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 290
 QY 1742 GCGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 1801
 DB 291 GCGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 350
 QY 1802 TGACGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 1861
 DB 351 GCGCGGCGGCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCT 401

QY 1862 CGTGAAGCGCTGCGGCTGACGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAGCT 1921
 DB 402 CGCGGCTGACATCCGCTG--ATCGGCGGCTGCGGAGTCTCTCGC--CGCGCAAGGT 458
 QY 1922 CGACCGGCTGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 1981
 DB 459 CAACGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 518
 QY 1982 GGTGCTGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCT 2041
 DB 519 GACCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCT 578
 QY 2042 CTACCTGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2101
 DB 579 GACCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCT 638
 QY 2102 CGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2161
 DB 639 GATGCGCTTCCACAGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGG 698
 QY 2162 GATCAAGTCATTCAGCTCCAGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGG 2221
 DB 699 CACCGCGGCGGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCT 758
 QY 2222 CT 2223
 DB 759 CT 760

RESULT 5
 us-11-143-980-1
 ; Sequence 1, Application US/11143980
 ; Publication No. US20050272133A1
 ; GENERAL INFORMATION:
 ; APPLICANT: He, Min
 ; APPLICANT: Huculi, John
 ; APPLICANT: Hattli, Bradley A.
 ; APPLICANT: Magenar, Melissa M.
 ; APPLICANT: Graziani, Edmund
 ; APPLICANT: Summers, Mia
 ; APPLICANT: Kulowski, Kerry
 ; APPLICANT: Pong, Kevin
 ; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
 ; FILE REFERENCE: AM-10142605
 ; CURRENT APPLICATION NUMBER: US/11/143,980
 ; PRIOR FILING DATE: 2005-06-03
 ; PRIOR APPLICATION NUMBER: US 60/664,483
 ; PRIOR FILING DATE: 2005-03-23
 ; PRIOR APPLICATION NUMBER: US 60/576,895
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: Patencin version 3.3
 ; SEQ ID NO 1
 ; LENGTH: 116856
 ; TYPE: DNA
 ; ORGANISM: Streptomyces sp.
 ; US-11-143-980-1
 Query Match 1.7%; Score 54.8; DB 7; Length 116856;
 Best Local Similarity 47.7%; Pred. No. 0.054;
 Matches 228; Conservative 0; Mismatches 242; Indels 8; Gaps 2;
 QY 1921 CGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1980
 DB 51048 GAGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 51107
 QY 1981 CGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 2038
 DB 51108 CGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 51167
 QY 2039 TCGCTACTGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2098

Db 51168 GCGCTTGGCCGAGGGGTAGACACACGAGGTCGAGGTGCGTGGCCGCTGTTCCGGG 51227
QY 2099 GCGCCCGAGAGACCGCAAGCCCAAGG-----CCTGTCGATTCCAGTGGATGGAAC 2152
Db 51228 GCGCCCGGGGTGATCTGCCCCACGCTAGCCCTTCCAGGCGGAGCCCTCTGGCTGGAGCG 51287
QY 2153 GCGCCCTCGATCAGTCCATGCTCCAGGAGCTCGGGGATTTCAGAGGGCCAAAGCG 2212
Db 51288 GCGCCACCGCGAGGCGCCCAACGAGCGGATCGACCGGAAATTCTGGCCCGCTTGAGCG 51347
QY 2213 GAGCGGATCTTCGCGCGGACACGCGCTGTCCGAGATTGTCGCCGCTCAAGAGGA 2272
Db 51348 CGAGGACCTTCGAGTGCCTGCGCGGACGCTGCGCTGACGCGGCTGCGGAAGT 51407
QY 2273 GGTGCTGGCCCGAGGAGTGGAGAAAGTCCGCAAGGACATTAGCGGTACTTAGG 2332
Db 51408 GCTGCGCGGCTGTCCCACTGGCGGCGCAAGCGCGTGAAGTCTCACATGACTCATG 51467
QY 2333 CTTTACTGACAGTAGAGCGCCGCTGCGCGGCGCGCGCGCGGCGGAGCC 2390
Db 51468 GCGTTACAGATCCGCTGAGAGCCGCTCACCCCGCGCCACTTACAGGACGAGCAC 51525

RESULT 6

US-11-075-185-36
; Sequence 36, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-36

Query Match 1.7%; Score 54.2; DB 7; Length 5679;
Best Local Similarity 42.8%; Pred. No. 0.024;
Matches 345; Conservative 0; Mismatches 453; Indels 9; Gaps 1;

QY 1646 GCGGGGCGGGGCGGGGCGGGTACGAGCCGACGCTGCCGCAACCAACGAGCTGCT 1705
Db 1356 GCGGACGTCATCATCGAGAGAGCGCGCGCGAGAGCGCGCGCGCAAGACGAGCGG 1415
QY 1706 GTGCGCGGACGAGCGGAGAGACCGGGCGCGCTTCGCGCAAGCTGTGTTGTACGCG 1765
Db 1416 AGCGGGGAGACAGAGCGCGGCGGGGTGTGCTGTGACGCTGCGCTCTCTGTG 1475
QY 1766 GCGCAACAACCGGCTGAGCTTCGCGGCTCGGCTATGACGAGCGACGAACTTCGCGG 1825
Db 1476 GGGTGGGACGAGCGGCGCTGCGCGCCAGGCGGGGCGGCTGCGCGAGCACTTCGCGG 1535
QY 1826 CAACGCGGACGCTGCTCTTTACGCGGCGGGCGGGGTGAAGCGCTGCGCTGCAAGG 1885
Db 1536 CCACCGGACGTCGGGTGCTGAGGTCGCGGGGCTGCGCAAGCGCGACGACCACT 1595
QY 1886 TGCAGGCTCATGCGCGCGCGCTGAGCTACGCGCAACCGTGTGGGTGGGGCGCG 1945
Db 1596 GCGCAGCGGCTGCGCGCTGCGCTGCGCGCGGACGAGCGAAGAGACTGGGCGCGG 1655
QY 1946 CAGTCCCCCGAGTACTGCGGACCAAGTCGGGCTGCTGCGCTGCGCGCAACG 2005

Db 1656 CTTTGGCAGTTCGCGCGCGG-----GGCCCGGCGCCGACGAGCGCGCCGCTGAC 1706
QY 2006 GCGCGCGGCGCGCGCGGATAGCGCGGCGCAATTCCTACTGTCGAGGAGCGGAGG 2065
Db 1707 GCGCGCGGAGCGCGCGCGGCAAGTGGGTGCTCTTACCGGCGGCGAGCGAGCGG 1766
QY 2066 CTTGCGCGGAGCGCTGCGCGCTGCGCGCGCGCGCGCGGAGGACGCAAGCGCAAGGA 2125
Db 1767 GCGCGGATAGGAGGCGCGCTGTATGCAACCAACCGCTTTCGCGCGCGCGCTGACG 1826
QY 2126 CTTGTCGATTCGAGCTGATTCGAGAGCGCCCTCTCTGATCAAGTTCATGCTCAGGA 2185
Db 1827 GCGCTGCGGAGCTTCAGCGGACGCTGACCGGCGCGCGCTGAGCTCTTTCGCGGA 1886
QY 2186 CTGCGGAGTTTACGAGAGCGCAAGGCGGAGCGGATTCGCGCGGCGGACAGCGCGGTC 2245
Db 1887 GCGCGGCTCGAGGCGCGCGGCTGTGACACGACAGCTGGGCTTCAGCGCGCTGTT 1946
QY 2246 CGAGATTTCGTCGCGCTCAAGAGCGAGTGTGGCCCAAGCGGAGCTGCGAGAACTG 2305
Db 1947 GCTCTGAGAGTGCAGCTTACCGGACAGTGGAAAGCTGGGGCTTGGCGCCGCGGCT 2006
QY 2306 GCGCAAGACATTAGCGGCTACTATGAGCTTCTACTGACAGCTTAGGCGCGCTGCGG 2365
Db 2007 GCTGCGGACAGCTCGCGGAGATGTCGCGCGCCACATCGCGGCTGCTGCACTTCA 2066
QY 2366 CCGCGCGCGCGCGCGCGGAGCGGAGCGGAGCGGCGGAGCGGCGGAGCGGCGGCGG 2425
Db 2067 CGAGCGCTCGCGCTGTGTGCGCGCGCGGAGCGGAGCGGCGGAGCGGCGGCGGCGG 2126
QY 2426 CTCGCCACACTCTCTTTCGCGGACCA 2452
Db 2127 GCGCATGAGCTTCATCGAGGCGACCGA 2153

RESULT 7

US-11-075-185-2/c
; Sequence 2, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 14172
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-2

Query Match 1.7%; Score 54.2; DB 7; Length 14172;
Best Local Similarity 45.2%; Pred. No. 0.034;
Matches 286; Conservative 0; Mismatches 338; Indels 9; Gaps 2;

QY 1519 GCTGACCCCTCGCCCAAGACTGCGCGGCTGCGAGATGTGCCGGGCGCGGCTACG 1578
Db 2636 GCGGAGACTGCTGATCGATCGATCGGCTGACCTTCTGTCTGAGGCGCTTCGCGG 2577
QY 1579 CCATGCGCGGCTCTTCTGTCGAGACAGATTGAGAACCTAGCCAAAGCGCGCTTC 1638
Db 2576 GCTGCGCTTCTGCGCGGCGGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 2517
QY 1639 ACCGGGCGGCGGCGGCGCGCGCGGCGGCTGAGACCGGAGCGTGCAGCAACG 1698


```

Db 30482 CGCGCGGCGCTCGCGCGTGGCGCACCGAAAGCTCCGGCTCCGTCGCCGCGGCGCAGCGCCG 30423
Qy 1587 GGGCTTTTCTGTGAGAGACAGAGTTGTGTAGCACTAGCCAGAGCCCGCTTTCACCCGGGC 1646
Db 30422 ACGCAGAGAGCGCCCCGTCACAGCGGCTGCGCGCCCCACGTCCAGACCCCGAGCCGCGCGTTC 30363
Qy 1647 GCGG--GCGCGGGGCCCCGGGCGGGGTACAGGACCGCGACGTCGCGGCAACCAACGGGCTGC 1704
Db 30362 GGGGTACTCGGCGCGCGTGGCCCGCAGACAGGCCCCAGACGGCCGCGCGGCCCGCGCGTTC 30303
Qy 1705 TGTGCGCGCAGACAGAGCCGAGAGACCCGGGACGCGCCCTTCGCGGCAACGCTGGTTGTGACGC 1764
Db 30302 CAGCGCCCTCGTCGCGCCCGGGTCGCGCACCGGCGACGCGGTCAACACCAAGAGGCTCGTCCG 30243
Qy 1765 CGCGCCAAACACCGGCTGGACTTTCGCGGCGCTTCGCGCTATATGACAGGCGCAAGGACTTTCGCG 1824
Db 30242 C-GCCAGGTGGGCGCGCGCGAGCCACCCCTTGACACAGGGCGACAGCGCGCGCGCGCACCGCG 30184
Qy 1825 GCAACGCGGCGCAGCGTGTCTTTACGCGGCGGGCGGGCGTGAAGGCGCTG--CCGCTGCA 1882
Db 30183 CGTGCAGGCGCGCGCACAGACGCGGCTGCAATCGCCCGCTCAGAGTCCAGACCAAGCCGCG 30124
Qy 1883 GGGTGCAGGCTGTGACTGAGCCGCGCCGCTC---GGCTACTAGCGCGAACCCGTCGGGCTGG 1938
Db 30123 GCGGCGGCGGCGCGCGCGCTCGCGCACGCGCGCTGAGAGAGCTCGGACAGCGCGCGCGCGCG 30064
Qy 1939 GCGCCCGGAGTCCCCCGGAGTACTGCGGCGCACCAATCGGCGTGGTGGCTGCGCTGCGC 1998
Db 30063 GCAGCGCCAGCGCGCGCGCGCGCACCGTACACCGCGCGACAGAGTGCTCTTCGCGGG 30004
Qy 1999 CCAACAGCGCGCGCGCGCGCGCGCGCGATAGCGCGCGCGCGCAATCCCTTACGTCGGGAGAG 2058
Db 30003 CCGACGTGCGCGCTTCAGAGCGCCACGTCCGTTCACGAGACCGCGTACATCGGTCCGCGC 29944
Qy 2059 CCGAGGCGCTTGGCGCGCGAGCGCTCGCGGCTGCGCGCGCGCGCGCGCGCGAGAGCGCAAC 2118
Db 29943 GGGTTCGGGCGGAGACCGGCTCTGCGCGCGCGCTCGCGCGCGCGCACCGCGCGGTGTGCCA 29884
Qy 2119 CCAAGAGCTGTCCGATTTCCAGTGTAGTCAAGAGCGCCCTTCGTGATCAAGTCAATGACT 2178
Db 29883 CGGTTTCGTCGTGGCGCTCGCGCGCAGACAGACGACGACACGCGCGCGCGCGCCCTCCG 29824
Qy 2179 CCAGCGACTCGGGGATTTAGAGAGAGGCGCAAGCGAGAGCGGAGTCTGCGCGCGAGACGC 2238
Db 29823 CCTGCTCCAGGCGCATGCGCGCGCGCGAGCTTGCTGCGCGCGGTGCTGATACAGGCGCAGT 29764
Qy 2239 CCGTGTCCGAGA 2250
Db 29763 CCGTCCAGCGCA 29752

RESULT 10
US-11-110-082-20
; Sequence 20, Application US/11110082
; Publication No. US20050266558A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Hall, Claire
; APPLICANT: Norries, Michael Geoffrey
; APPLICANT: Saulebury, Keith Martin
; TITLE OF INVENTION: Compositions Isolated from Forage
; FILE REFERENCE: 1100.1074UC1
; CURRENT APPLICATION NUMBER: US/11/110,082
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 60/563,723
; PRIOR FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 10/655,799
; PRIOR FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/408,762
; PRIOR FILING DATE: 2002-09-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 20
; LENGTH: 2302
; TYPE: DNA
; ORGANISM: Festuca arundinacea
US-11-110-082-20

Query Match      1.6%; Score 52.2; DB 7; Length 2302;
Best Local Similarity 44.3%; Pred. No. 0.047;
Matches 213; Conservative 0; Mismatches 268; Indels 0; Gaps 0;

QY      1706 GTTCGCCGACAGAGCCGAGAGACCAGGCGCCCTTCGCGCAACGCTGGTTGTGACGCC 1765
DB      585 GCCGGCGCAGCTGTTCGCCGACCTGTCTGCGCACTCCCGCGGAGACCCCAAGCGGATGCG 644
QY      1766 GGCCTAACCAACCGGCTGAGACTTCCGGGCTTCGGGCTATGACACGGCCACAGACTTTCGGGG 1825
DB      645 TACCGCGCGCGGAGAGACGCTTCCTTCATCATGTCATTCCTTCCTTCGGCGGCTGCGT 704
QY      1826 CAACGCGGCGACGCGTCTCTCTTACGCGGCGCGCGGCGGTGAAGGCGCTGCGCTGCAGAGC 1885
DB      705 GGTGAGAGCGCTCTCCGCGCGGCGCGCGAGGCGCAACGCCATCGCGCTGCGCGGTCTGTGT 764
QY      1886 TGCAGGCTTGCACTGAGCGCGCGCGCTCGGCTTACGACCGACCCGTGCGGCTGCGGCGCGCG 1945
DB      765 GGCAGCAGACGAGGAGGAGGAGATCCGGGCTGTGACAGCGCGCTGTGGGTGTGCGCGAGAGC 824
QY      1946 CAGTCCCCCGCAGTACTGCGGACACCAAGTGGGCTCGGTGTGCTGCTGCTGAGCCCAACAG 2005
DB      825 CGTGCAGCAGGAGAACTTCTTCGCGCGCGCGAGGCGGTGTGAGAGAGATACCTTGTGAGC 884
QY      2006 GCGCGCGGCGCGCGCGCGCATGCGCGCGCGGCGCAATCCCTACCTGCGGAGAGGCGCGAGG 2065
DB      885 GGCCTTCCAGGCGGCGCGCATGCGCAAGTGCAGGCTTACTTTCGAGAGCCCTTGCGCG 944
QY      2066 CCTGCGCGCGGAGCGCTGCGCGCTGCGCGCGCGCGCGCGGAGAGCGCAAGCCCAAGG 2125
DB      945 CCGCGCTTCCGCTTCCGCTTCCGCGCGCGCGCGAGCGCAGGTCCACCTGACGCGCGCTTGCGGA 1004
QY      2126 CCTGTCCGATTCCAGCTGATGATGACAGACGCCCTCTCGATCAAGTTCATCGACTCCAGGGA 2185
DB      1005 CCTCCTCCACGCGCATTTCTACGAGTCTGCCCCCTACCTCAAGTTGCGCCACTTCACCGC 1064
QY      2186 C 2186
DB      1065 C 1065

RESULT 11
US-11-045-802-25/c
; Sequence 25, Application US/11045802
; Publication No. US20050257289A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Helentjaris, Tim
; APPLICANT: Lowe, Keith
; APPLICANT: Shen, Bo
; APPLICANT: Tarczyński, Mitchell
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: Ap2 Domain Transcription Factor ODP2 (Ovule Development Protein 2
; FILE REFERENCE: 035718/286074
; CURRENT APPLICATION NUMBER: US/11/045,802
; PRIORITY FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 60/541,122
; PRIORITY FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 2079
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc_feature

```


Db 6054 GGGCGCGGCGCGCGCGCGCGAGTGCAGGCCCGCCAGGACGACGCGGGAACCTCCG 5995b
 QY 1934 CTGGGGCGCCCGCAGTCCCGCCAGTACTGCGGCGACCAAG--TCGGGCTCGGTGTGTCGC 1991b
 Db 5994 CCGCGCGATTGGCCCGCGCGCGCGCGCCCGCCAGCACACCAATTTGGCCCGCGCTGACG 5953b
 QY 1992 TGCTGGCCCAACAGCGCGCGCGCGCGCGCGCATGGCGGCGCAATTCCTACTGTGGGC 2051b
 Db 5934 CATTGTTCGTGACTGCGCGCGCGGCGGCGGCGACCGCTGGACCCCGGAAAACTTCCGA 5875b
 QY 2052 GAGGAGGCGGAGGGCTGTGCGCGCAGGCGCTTCGCGCTGCGCGCCCGCGCGCCGCGCA 2107b
 Db 5874 GAAACTTGAAGGAGGCGCGCGCGGAGAAACGAGAGGAGGCGGCGCGCGCTGTGGCGA 5819b

```

RESULT 14
US-10-432-483-22
Sequence 22, Application US/10432483
Publication No. US20050260699A1
GENERAL INFORMATION:
APPLICANT: desouza, Mervyn L.
APPLICANT: Jessen, Holly
APPLICANT: Schroeder, William A.
APPLICANT: Gokarn, Ravi R.
TITLE OF INVENTION: CAROTENOID BIOSYNTHESIS
FILE REFERENCE: 12799-0020U1
CURRENT APPLICATION NUMBER: US/10/432,483
CURRENT FILING DATE: 2003-11-10
PRIOR APPLICATION NUMBER: PCT/US01/43906
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/252,749
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 996
TYPE: DNA
ORGANISM: Micrococcus luteus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(993)
US-10-432-483-22

```

	Query Match	1.6%	Score 51.6;	DB 6,	Length 996;
	Best Local Similarity	51.3%;	Pred. No. 0.046;		
	Matches	120;	Conservative	0;	Mismatches 114; Indels 0; Gaps 0;
Qy	1511	CATGAGACGGCTGACCCCTTCGCGCCCAACGACTCGCGCGCTCGACGAGTGTGCCCGGAGC	1570		
Db	489	CATGGAATGACGCCCCGCGCACCCGCGCCCAACCCCGAGACAGCGGAGATGTCTGCAGCAC	548		
Qy	1571	CCGCTAACCCATGCGCGGCTCTTTTCCTGACAGACAGATTCTGTAGCAACTTAGCCCAAGC	1630		
Db	549	GCGCCGCGGCTGGGTGCGCGGCTTCCAAGAAGTCACTTCTTGCGGGAATCTGGGCGGAA	608		
Qy	1631	CCGCTTTCACCCCGGCGCGGCGCGCGGCGCCGCGGCGCGGATAACGACCGCAGCGTGCACCA	1690		
Db	609	CCAGGACCAGCTCGGACGCACTTAATTCCTCCGCGCGGAGCCCTCCACCTGACGAAAC	668		
Qy	1631	CACCAACGGGCTGTCTGTGCGCGGCAACAGACGAGCCGAGAACCCCGGCGCGCCCTTGCC	1744		
Db	669	CCGCAAGGGGCTGTCTGTGCGGAGACTTGGCGCGGACCTTGACCGGCGCGGCGTGGCC	722		

RESULT 15
US-10-432-483-49
; Sequence 49, Application US/10432483-49
; Publication No. US20050260699A1
; GENERAL INFORMATION:
; APPLICANT: desouza, Mervyn L.
; APPLICANT: Jessen, Holly
; APPLICANT: Schroeder, William A.

```

; APPLICATION: Gokarn, Ravi R.
; TITLE OR INVENTION: CAROTENOID BIOSYNTHESIS
; FILE REFERENCE: 12799-002US1
; CURRENT APPLICATION NUMBER: US/10/432, 483
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: PCT/US01/43906
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/252, 749
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; TYPE: DNA
; LENGTH: 6941
; ORGANISM: Micrococcus luteus
US-10-432-483-49

Query Match      1.6%; Score 51.6; DB 6; Length 6941;
Best Local Similarity 51.3%; Pred. No.0.095;
Matches 120; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

1511 CATGAGCCGCGCTACACCCCTTCGCCCAACGACTGCGCCGCCCTCGAGAGTTCGCCCGGGGCG 1570
Db      2276 CATGAGACATGCGCCGCGACCCGCGCCGACGCCCGCGAAGCGGAGAGATGCTGCCCGCCAC 2335
Qy      1571 CCGGTACGCCCATGCGCGCGAGCTTCTTCCTCGACAGACAGAGTTCGTGAGCAACTACGCCAAGCG 1630
Db      2336 GCGCCGCGCGGTGGGTGCGCGGTTCCAGAAAGTCACTTCTCTGGGGATCTCGCGCGGGA 2395
Qy      1631 CCGGTTCCACCCGGGCGCGGGCGCGGGGCCCGGGCGGAGTACGAGCCGACGCGTCCGCA 1690
Db      2396 CCACGACAGAGCTCGGAGCGACACTTACTTCCCGGCGCGGACCCCTCCACCTGACGAGAC 2455
Qy      1691 CACCAACGGCGTGTGTCGTCCGCGACAGCGCCGAGAGACCCGGGCGCGCCCTTCGCC 1744
Db      2456 CCGAAGGCGGTGTGCTGTCGCGACCTTCGGCGGACCTTGACCGCGGCGCTGCC 2509

```

Search completed: December 20, 2005, 16:52:30
Job time : 352 secs

```

1  APPLICANT: Gokarn Ravi R.
2  TITLE OF INVENTION: CAROTENOID BIOSYNTHESIS
3  FILE REFERENCE: 12799-002051
4  CURRENT APPLICATION NUMBER: US/10/432,483
5  CURRENT FILING DATE: 2003-11-10
6  PRIOR APPLICATION NUMBER: PCT/US01/43906
7  PRIOR FILING DATE: 2001-11-21
8  PRIOR APPLICATION NUMBER: US 60/252,749
9  PRIOR FILING DATE: 2000-11-22
10 NUMBER OF SEQ ID NOS: 49
11 SOFTWARE: FaastSeq for Windows Version 4.0
12 SEQ ID NO 49
13 LENGTH: 6941
14 TYPE: DNA
15 ORGANISM: Micrococcus luteus
16 US-10-432-483-49

```

Match	1.6%	Score 51.6	DB 6	Length 6941
Local Similarity	51.3%	Pred. No. 0	0.095	
s 120	Conservative	0	Mismatches 114	Indels 0
			Gaps	0
1511	CATGAGACGGCTGACCCCTTGCCCAACGACTCGCCGCTCGCAGATGTGCCCGGAGC	1570		
2276	CATGAGCATGCGCCGACCCCGCCAGACCCCGGAGCAGCGGAGAGTCTGCGCGCAC	2335		
1571	CGCTAACGCATAGCGCGCTCTTTCGACGAGACGAGTTCGTAGCAACTACGCCAGGC	1630		
2336	GGCGCGCGGCTGGGTGCGGGTTCAGAAAGTCCAACTTCTCGGAGATCTGCGCGGAA	2395		
1631	CGGCTTCCACCCCGGCGCGGAGCGCGGAGCCCGGAGCGGGTACGAGCCGACGTCGCGCA	1690		
2396	CCACGACGACGCTCGGACGCACTTCTTCCCGGCGGCGGACCCCTCCACCTGACGAGAC	2455		
1691	CACCAACGGGCTGCTGTGTGCGCGACGACGAGCGGAGACCCGGGCGCGCCTCGCC	1744		
2456	CCGCAAGCGGCTGCTGTCTCGCGGACCTTCGGGCGCGGACCTGAGACCGCGCGTGTCC	2509		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 17:22:49 ; Search time 347 Seconds
(without alignments)
4500.606 Million cell updates/sec

Title: US-10-649-400-3
Perfect score: 3186
Sequence: 1 caggtgattacctaataa.....caccacaacaccttaag 3186

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 4161431 seqs, 245089505 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8322862

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications_NA_New.*
1: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	0.8	169725	7 US-11-121-086-63	Sequence 63, Appl
2	25	0.8	169725	7 US-11-121-086-63	Sequence 168788
3	25	0.8	169725	7 US-11-121-086-63	Sequence 168790
4	25	0.8	169725	7 US-11-121-086-63	Sequence 168791
5	25	0.8	169725	7 US-11-121-086-63	Sequence 168792
6	25	0.8	169725	7 US-11-121-086-63	Sequence 168793
7	25	0.8	169725	7 US-11-121-086-63	Sequence 168794
8	25	0.8	169725	7 US-11-121-086-63	Sequence 168795
9	25	0.8	169725	7 US-11-121-086-63	Sequence 168796
10	25	0.8	169725	7 US-11-121-086-63	Sequence 168797
11	25	0.8	169725	7 US-11-121-086-63	Sequence 168798
12	25	0.8	169725	7 US-11-121-086-63	Sequence 168799
13	25	0.8	169725	7 US-11-121-086-63	Sequence 168800
14	25	0.8	169725	7 US-11-121-086-63	Sequence 168801
15	25	0.8	169725	7 US-11-121-086-63	Sequence 168802
16	25	0.8	169725	7 US-11-121-086-63	Sequence 168803
17	25	0.8	169725	7 US-11-121-086-63	Sequence 168804
18	25	0.8	169725	7 US-11-121-086-63	Sequence 168805
19	25	0.8	169725	7 US-11-121-086-63	Sequence 168806
20	25	0.8	169725	7 US-11-121-086-63	Sequence 168807
21	25	0.8	169725	7 US-11-121-086-63	Sequence 168808
22	25	0.8	169725	7 US-11-121-086-63	Sequence 168809
23	25	0.8	169725	7 US-11-121-086-63	Sequence 168810

24	0.8	51917	6	US-10-995-561-13338	Sequence 13338, A
25	0.8	110847	7	US-11-121-086-63	Sequence 11, Appl
26	0.8	119160	7	US-11-121-086-63	Sequence 12, Appl
27	0.8	126552	7	US-11-121-086-63	Sequence 1, Appl
28	0.8	162085	7	US-11-121-086-63	Sequence 7, Appl
29	0.8	165156	6	US-10-995-561-13304	Sequence 13, Appl
30	0.8	169495	7	US-11-121-086-63	Sequence 61, Appl
31	0.8	169495	7	US-11-121-086-63	Sequence 62, Appl
32	0.8	285300	6	US-10-857-780-6	Sequence 6, Appl
33	0.8	305312	6	US-10-995-561-13236	Sequence 13236, A
34	0.8	1125000	6	US-10-995-561-13286	Sequence 13286, A
35	0.7	25	7	US-11-121-086-63	Sequence 168787, A
36	0.7	201	6	US-10-995-561-23615	Sequence 23615, A
37	0.7	201	6	US-10-995-561-23773	Sequence 23773, A
38	0.7	201	6	US-10-995-561-23802	Sequence 23802, A
39	0.7	201	6	US-10-995-561-23805	Sequence 23805, A
40	0.7	201	6	US-10-995-561-30516	Sequence 30516, A
41	0.7	201	6	US-10-995-561-43974	Sequence 43974, A
42	0.7	201	6	US-10-995-561-58486	Sequence 58486, A
43	0.7	201	6	US-10-995-561-58499	Sequence 58499, A
44	0.7	201	6	US-10-995-561-58501	Sequence 58501, A
45	0.7	201	6	US-10-995-561-58502	Sequence 58502, A

ALIGNMENTS

RESULT 1
US-11-121-086-63
; Sequence 63, Application US/11121086
; Publication No. US2005026459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121, 086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567, 570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 63
; LENGTH: 169725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-63

Query Match 0.8%; Score 26; DB 7; Length 169725;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGGATGATGATGATGATG 2841
DB 120068 GAGTGGATGATGATGATG 120093

RESULT 2
US-11-121-086-63
; Sequence 168788, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121, 849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567, 949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 168788

LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-11-121-849-168788

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2933 GTCAAGCACAAACCAAGTTCATGC 2957
DB 1 GTCAAGCACAAACCAAGTTCATGC 25

RESULT 3
US-11-121-849-168790
Sequence 168790, Application US/11121849
Publication No. US20050272080A1
GENERAL INFORMATION:
APPLICANT: John Palma

TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
FILE REFERENCE: 3684.1
CURRENT APPLICATION NUMBER: US/11/121,849
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/567,949
PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 673904
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 168790
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-11-121-849-168790

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3040 AGATAATATTCTAAGAGGCTCTAGA 3064
DB 1 AGATAATATTCTAAGAGGCTCTAGA 25

RESULT 4
US-11-121-849-168791
Sequence 168791, Application US/11121849
Publication No. US20050272080A1
GENERAL INFORMATION:
APPLICANT: John Palma

TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
FILE REFERENCE: 3684.1
CURRENT APPLICATION NUMBER: US/11/121,849
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/567,949
PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 673904
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 168791
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-11-121-849-168791

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 ATACTCAGTAGAGGAGGTTTCCCA 3097
DB 1 ATACTCAGTAGAGGAGGTTTCCCA 25

RESULT 5
US-11-121-849-168792
Sequence 168792, Application US/11121849
Publication No. US20050272080A1
GENERAL INFORMATION:
APPLICANT: John Palma

TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
FILE REFERENCE: 3684.1
CURRENT APPLICATION NUMBER: US/11/121,849
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/567,949
PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 673904
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 168792
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-11-121-849-168792

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3084 TGATGGGTTTCCCACTTCTCTCA 3108
DB 1 TGATGGGTTTCCCACTTCTCTCA 25

RESULT 6
US-11-121-849-168793
Sequence 168793, Application US/11121849
Publication No. US20050272080A1
GENERAL INFORMATION:
APPLICANT: John Palma

TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
FILE REFERENCE: 3684.1
CURRENT APPLICATION NUMBER: US/11/121,849
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/567,949
PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 673904
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 168793
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-11-121-849-168793

Query Match
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3121 AATAATTAATAGGAGGCTCTATGCA 3145
DB 1 AATAATTAATAGGAGGCTCTATGCA 25

RESULT 7
US-11-121-849-168794
Sequence 168794, Application US/11121849
Publication No. US20050272080A1
GENERAL INFORMATION:
APPLICANT: John Palma

TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
FILE REFERENCE: 3684.1
CURRENT APPLICATION NUMBER: US/11/121,849
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/567,949

PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 673504
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 168794
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-11-121-849-168794

Query Match 0.8%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3132 TGTGCTATGACACAAATAGCT 3156
DB 1 TGTGCTATGACACAAATAGCT 25

RESULT 8
US-10-995-561-15048/c
Sequence 15048, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15048
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-15048

Query Match 0.8%; Score 25; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2817 AGTGATGATGATGATGATGATG 2841
DB 31 AGTGATGATGATGATGATGATG 7

RESULT 9
US-10-750-185-26071
Sequence 26071, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMT GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIN version 3.1
SEQ ID NO 26071
LENGTH: 2095
TYPE: DNA
ORGANISM: Bovine 1986680854269
US-10-750-185-26071

Query Match 0.8%; Score 25; DB 6; Length 2095;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATGATG 2840
DB 344 GAGTGATGATGATGATGATGATG 368

RESULT 10
US-10-995-561-13198/c
Sequence 13198, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13198
LENGTH: 209822
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-13198

Query Match 0.8%; Score 25; DB 6; Length 209822;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2817 AGTGATGATGATGATGATGATG 2841
DB 51465 AGTGATGATGATGATGATGATG 51441

RESULT 11
US-11-121-849-168785
Sequence 168785, Application US/11121849
Publication No. US20050272080A1
GENERAL INFORMATION:
APPLICANT: John Palma
TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded
FILE REFERENCE: 3684.1
CURRENT APPLICATION NUMBER: US/11/121,849
CURRENT FILING DATE: 2005-05-03
PRIOR APPLICATION NUMBER: 60/567,949
PRIOR FILING DATE: 2004-05-03
NUMBER OF SEQ ID NOS: 673904
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 168785
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-11-121-849-168785

Query Match 0.8%; Score 24; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2870 AAGCTGTGAATGATGATGATGAT 2893
DB 2 AAGCTGTGAATGATGATGATGAT 25

RESULT 12
US-10-995-561-45127
Sequence 45127, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:

```

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45127
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-45127

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGGATGGATGGATGGATGGATGG 2841
DB      137 GTGGATGGATGGATGGATGGATGG 160

RESULT 13
US-10-995-561-50737
; Sequence 50737, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50737
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50737

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGGATGGATGGATGGATGGATGG 2841
DB      101 GTGGATGGATGGATGGATGGATGG 124

RESULT 14
US-10-995-561-50739
; Sequence 50739, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50739
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50739
```

```

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGGATGGATGGATGGATGGATGG 2841
DB      100 GTGGATGGATGGATGGATGGATGG 123

RESULT 15
US-10-995-561-50759
; Sequence 50759, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50759
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50759

Query Match          0.8%; Score 24; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2818 GTGGATGGATGGATGGATGGATGG 2841
DB      99 GTGGATGGATGGATGGATGGATGG 122
```

Search completed: December 21, 2005, 02:00:48
Job time : 350 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 15:56:28 / Search time 15680 Seconds
(without alignments)
11549.951 Million cell updates/sec

Title: US-10-649-400-3
Perfect score: 3186
Sequence: 1 caggtgattatcctaataa.....cacccaacaccttaagg 3186

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size: 0

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sta:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2617	82.1	2910	6	CQ729238 Sequence
2	2259	70.9	2894	8	HSU49250 Human putat
3	1793	56.3	1793	6	CQ769652 Sequence
4	1793	56.3	1793	8	BC028285 Homo sapi
5	1694	53.2	183666	8	AC009487 Homo sapi
6	514	17.2	101040	14	AC165777 Bos lauru
7	514	16.1	514	6	AR655897 Sequence
8	384	12.1	388	6	BD024632 Sequence
9	384	12.1	388	6	AX885022 Sequence
10	380	11.9	641	8	HS438878 Homo sapi
11	367	11.5	689	8	HS434262 Homo sapi
12	343	10.8	793	8	HS4341671 Homo sapi
13	289	9.1	289	8	AY271043 Homo sapi
14	266	8.3	693	8	HS4341518 Homo sapi
15	235	7.4	235	6	BD041913 Sequence
16	235	7.4	235	6	AX906380 Sequence
17	206	6.5	3897	9	BC058399 Mus muscu
18	206	6.5	3992	9	BC052737 Mus muscu

C	19	206	6.5	176172	14	AC127930 Rattus no
C	20	206	6.5	209914	9	AL845291 Mouse DNA
C	21	155	3.7	3814	9	MMU49251 Mus musculu
C	22	117	4.9	701	8	HS4341512 Homo sapi
C	23	107	3.4	671	8	HS4341522 Homo sapi
C	24	89	2.8	4570	9	AB032374 Homo sapi
C	25	84	2.6	691	8	HS4339038 Homo sapi
C	26	78	2.4	101040	14	AC165777 Bos lauru
C	27	60	1.9	60	6	CQ536907 Sequence
C	28	51	1.6	690	8	HS4325162 Homo sapi
C	29	47	1.5	784	8	HS4327080 Homo sapi
C	30	33	1.0	741	8	HS4332554 Homo sapi
C	31	32	1.0	951	8	HS4325603 Homo sapi
C	32	29	0.9	174764	9	AC126257 Mus muscu
C	33	29	0.9	220887	9	AC124741 Mus muscu
C	34	29	0.9	224279	9	AC102135 Mus muscu
C	35	29	0.9	261825	9	AC146300 Mus muscu
C	36	28	0.9	188317	14	AC148957 Otlemur
C	37	28	0.9	216447	14	AC133122 Rattus no
C	38	27	0.8	60	6	BD043446 Sequence
C	39	27	0.8	60	6	AX907913 Sequence
C	40	27	0.8	532	2	AY700569 Schistoc
C	41	27	0.8	547	5	AF069393 Gallus ga
C	42	27	0.8	1246	5	AF033668 Gallus ga
C	43	27	0.8	1317	5	AY179342 Lampetra
C	44	27	0.8	2855	5	AB019787 Cynops py
C	45	27	0.8	30804	8	AC134338 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS CQ729238 2910 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15172 from Patent WO02068579.
ACCESSION CQ729238
VERSION CQ729238.1 GI:42300290

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof: WO 02068579-A 15172 06-SEP-2002;

JOURNAL

PE Corporation (NY) (US)
Location/Qualifiers
1..2910
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES

source
1..2910
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 82.1%; Score 2617; DB 6; Length 2910;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2907; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 279 GGGAGGTTCAGGTTCTAGAGCTATGACGCTGAGCACTGCTTCTCTTATCAG 338
1 GGGAGGTTCAGGTTCTAGAGCTATGACGCTGAGCACTGCTTCTCTTATCAG 60
QY 339 CTCTCCAAAGAAATTTCTCATGAGGACGACGATACCACTTACAGCGGAGTTC 398
61 CTCTCCAAAGAAATTTCTCATGAGGACGACGATACCACTTACAGCGGAGTTC 120
QY 399 GTCTTCAGGATCATCCATTAATCTCGACCACTGACCACTGAGAGAAAGTTACCTTG 458
121 GTCTTCAGGATCATCCATTAATCTCGACCACTGACCACTGAGAGAAAGTTACCTTG 180

OY	459	AAAAAAATTACAGGGGGATGACCAATAGTCAGATACAGCAATTTTCTGATCTCCAG	518
Db	181	AAAAAAATTACAGGGGGATGACCAATAGTCAGATACAGCAATTTTCTGATCTCCAG	240
OY	519	GACTCACCAGGGAGCGTCCAGAGAACTAACTCTCTCTGTGCTTGAGACGGGGTCTGAG	578
Db	241	GACTCACCAGGGAGCGTCCAGAGAACTAACTCTCTCTGTGCTTGAGACGGGGTCTGAG	300
OY	579	CTTGCTCAAGTTTCGATGGCTCTGCTGCAGATGGCTACCTCTCTCTCACTCAGCCAG	638
Db	301	CTTGCTCAGAGTTTCGATGGCTCTGCTGCAGATGGCTACCTCTCTCTCACTCAGCCAG	360
OY	639	CCACAGCTGGGGGCCACTGCTCCAGATGCCAATGTTCCCGTACCCCGGACACAGACCG	698
Db	361	CCACAGCTGGGGGCCACTGCTCCAGATGCCAATGTTCCCGTACCCCGGACACAGACCG	420
OY	699	GCGACCCCGGCTTCTCCATCGGACGCCCTAGCCGCTACGATGGCCACACCCGATATC	758
Db	421	GCGACCCCGGCTTCTCCATCGGACGCCCTAGCCGCTACGATGGCCACACCCGATATC	480
OY	759	ACCAACGAGAGCTTACAAACGCTCTCTGTTCCACTCTTGCCGCGAGAGATACCCCAAGGC	818
Db	481	ACCAACGAGAGCTTACAAACGCTCTCTGTTCCACTCTTGCCGCGAGAGATACCCCAAGGC	540
OY	819	GAGTACCCCTTACCCACAGACAGTACCGGCACTCTTACCAAGAGCTCCGTTCTACAGTTC	878
Db	541	GAGTACCCCTTACCCACAGACAGTACCGGCACTCTTACCAAGAGCTCCGTTCTACAGTTC	600
OY	879	TCCTCCACCCAGCCGGGGCTGGTGCCCGGCAAAAGCAAGGTGTACTGTGCAACAGGCC	938
Db	601	TCCTCCACCCAGCCGGGGCTGGTGCCCGGCAAAAGCAAGGTGTACTGTGCAACAGGCC	660
OY	939	CTTTGGCTGAATTTACCCGGCACAAACGGAGATGATCATCACCAACAGGAAAGGCGC	998
Db	661	CTTTGGCTGAATTTACCCGGCACAAACGGAGATGATCATCACCAACAGGAAAGGCGC	720
OY	999	ATGTTTCTTTTAAAGTTTAAACATTTCTGGTCTGATCCACGGCTCATTTACATATTT	1058
Db	721	ATGTTTCTTTTAAAGTTTAAACATTTCTGGTCTGATCCACGGCTCATTTACATATTT	780
OY	1059	TTTGTGATGTGATTTTGGCGGATCCCAATCATCTGGAGTTTCAAGAGGCAATGGTT	1118
Db	781	TTTGTGATGTGATTTTGGCGGATCCCAATCATCTGGAGTTTCAAGAGGCAATGGTT	840
OY	1119	CTTTCGGCAAAAGCGAGACCAATGTGAAAGAAATCGGCTCTATATCATCTCGGATTC	1178
Db	841	CTTTCGGCAAAAGCGAGACCAATGTGAAAGAAATCGGCTCTATATCATCTCGGATTC	900
OY	1179	CCCAACACTGGGGCTCACTGATGCGCAAG-AAATCTCTTTGGAAATTTAAACTTTAC	1237
Db	901	CCCAACACTGGGGCTCACTGATGCGCAAGAAATCTCTTTGGAAATTTAAACTTTAC	960
OY	1238	GAACAACAAA-GGAGCTTCAAAATAAATGGGCAATGATGGTTTTTCAAGTCTTTGACA	1296
Db	961	GAACAACAAAAGGAGCTTCAAAATAAATGGGCAATGATGGTTTTTCAAGTCTTTGACA	1020
OY	1297	AGTACCAAGCCCGCCTGATGTGTGAAAGTAAAGAGACGGACCGAGGACACTAGCC	1356
Db	1021	AGTACCAAGCCCGCCTGATGTGTGAAAGTAAAGAGACGGACCGAGGACACTAGCC	1080
OY	1357	AGCCCGGCGCTGAGAGCTTCACTTTCCCTTGAACTCACTGATTCAGCCGCTACCGGCT	1416
Db	1081	AGCCCGGCGCTGAGAGCTTCACTTTCCCTTGAACTCACTGATTCAGCCGCTACCGGCT	1140
OY	1417	ACCAACAACGAGATTTTACACAACCTGAAATATGATCAACCCCTTTGCAAAAGATTC	1476
Db	1141	ACCAACAACGAGATTTTACACAACCTGAAATATGATCAACCCCTTTGCAAAAGATTC	1200
OY	1477	GGGATTAATTATGACAGATCTACACCGGCTGTGACATGAGCCGCTGACCCCTTGCCCA	1536
Db	1201	GGGATTAATTATGACAGATCTACACCGGCTGTGACATGAGCCGCTGACCCCTTGCCCA	1260

[illegible]

QY	939	CTTTGGCTGA	AAATTTAC	CCGGCAGC	CAACCGA	AGATGAT	CATCAC	CAACAC	AGGAGG	CGC	998	
Db	661	CTTTGGCTGA	AAATTTAC	CCGGCAGC	CAACCGA	AGATGAT	CATCAC	CAACAC	AGGAGG	CGC	720	
QY	999	ATGTTTCTTT	TTTAAGTTT	TAACTTT	CTGGTCT	CGATCC	CAGGCT	GCATTA	CAATTT	1058		
Db	721	ATGTTTCTTT	TTTAAGTTT	TAACTTT	CTGGTCT	CGATCC	CAGGCT	GCATTA	CAATTT	780		
QY	1059	TTTGTGAT	TGTGATTT	TGGCGAT	CCCATCA	CTGGAG	GTTC	CAAGAG	GCMAAT	TGGGTT	1118	
Db	781	TTTGTGAT	TGTGATTT	TGGCGAT	CCCATCA	CTGGAG	GTTC	CAAGAG	GCMAAT	TGGGTT	840	
QY	1119	CCTTGGCGCA	AGCGGAG	CACCAAT	GTGCAAG	AAATCGGG	CTPAT	ATGCAT	CCGGAT	TCC	1178	
Db	841	CCTTGGCGCA	AGCGGAG	CACCAAT	GTGCAAG	AAATCGGG	CTPAT	ATGCAT	CCGGAT	TCC	900	
QY	1179	CCCAACACT	GTGGGCTCA	CTGTGAT	GGCCAG	AAATCTCT	TTTGG	AAAAATT	AAAACTT	TACG	1238	
Db	901	CCCAACACT	GTGGGCTCA	CTGTGAT	GGCCAG	AAATCTCT	TTTGG	AAAAATT	AAAACTT	TACG	960	
QY	1239	AACAAACA	AGAGGCTT	CAATTA	PAACAAT	GGGAGAT	TGGTGT	TTTAAC	GTCTT	GGACAAG	1298	
Db	961	AACAAACA	AGAGGCTT	CAATTA	PAACAAT	GGGAGAT	TGGTGT	TTTAAC	GTCTT	GGACAAG	1020	
QY	1299	TACCAAGCCC	CGCTCAT	GTGGTGG	AAATGTAA	CAGACAG	CGCAG	CAAGAG	CACTA	AGCCAG	1358	
Db	1021	TACCAAGCCC	CGCTCAT	GTGGTGG	AAATGTAA	CAGACAG	CGCAG	CAAGAG	CACTA	AGCCAG	1080	
QY	1359	CCCGGCGCG	TGCAGAG	CGTTCA	CTTTCC	CTGAGACT	CAGTTCA	TGATGCC	CTGAC	CGGCTAC	1418	
Db	1081	CCCGGCGCG	TGCAGAG	CGTTCA	CTTTCC	CTGAGACT	CAGTTCA	TGATGCC	CTGAC	CGGCTAC	1140	
QY	1419	CAGAACAC	AGGATTT	TACACA	CACTG	AAATAT	ATATCA	CAACCT	TTTGG	CAAAAGATTT	1478	
Db	1141	CAGAACAC	AGGATTT	TACACA	CACTG	AAATAT	ATATCA	CAACCT	TTTGG	CAAAAGATTT	1200	
QY	1479	GATAATTA	TATGAC	AGATCT	TACAC	CGGGCT	TGTGAT	GACGCG	CTTGAC	CCCCCT	TGCGCAAC	1538
Db	1201	GATAATTA	TATGAC	AGATCT	TACAC	CGGGCT	TGTGAT	GACGCG	CTTGAC	CCCCCT	TGCGCAAC	1260
QY	1539	GACTGGCG	CGCGTGC	CAATG	TGTGCG	GGGGCG	CGCTPA	CGCAT	TGGCGG	CTCTTT	CTGT	1598
Db	1261	GACTGGCG	CGCGTGC	CAATG	TGTGCG	GGGGCG	CGCTPA	CGCAT	TGGCGG	CTCTTT	CTGT	1320
QY	1599	CAGGACCA	GATTCGT	TAGGACA	CTTAC	CGCAGG	CCCGCTT	CCA	CCGCGG	CGCGG	CGCGGC	1658
Db	1321	CAGGACCA	GATTCGT	TAGGACA	CTTAC	CGCAGG	CCCGCTT	CCA	CCGCGG	CGCGG	CGCGGC	1380
QY	1659	CCCCGGG	CGGGAT	ACCGA	CCGCA	GGGCTG	CGCCAC	CAAC	CGGGCT	GTCTG	CGCCGCA	1718
Db	1381	CCCCGGG	CGGGAT	ACCGA	CCGCA	GGGCTG	CGCCAC	CAAC	CGGGCT	GTCTG	CGCCGCA	1440
QY	1719	GCCGAGAG	ACCCGGG	GCGCC	CTCG	CGGCA	ACGCTG	TTTGT	GTGAC	CGCGG	CAACAA	1778
Db	1441	GCCGAGAG	ACCCGGG	GCGCC	CTCG	CGGCA	ACGCTG	TTTGT	GTGAC	CGCGG	CAACAA	1500
QY	1779	CTGACCTT	CGCGGCT	CGGCTAT	GAACA	CGGAC	CGAC	CGA	CTTGTG	GGGCAAC	CGCGGCA	1838
Db	1501	CTGACCTT	CGCGGCT	CGGCTAT	GAACA	CGGAC	CGAC	CGA	CTTGTG	GGGCAAC	CGCGGCA	1560
QY	1839	CTGCTCT	CTTAC	GGCGG	CGCGG	CGGTGA	AGGCG	CTG	CGGCTG	CAAGG	CTGCA	1898
Db	1561	CTGCTCT	CTTAC	GGCGG	CGCGG	CGGTGA	AGGCG	CTG	CGGCTG	CAAGG	CTGCA	1620
QY	1899	GGCGCG	CGCGCT	CGGCTACT	ACGAC	CGGCTG	CGGCTG	GGGCG	CGCCG	CAAGT	CCCCCG	1958
Db	1621	GGCGCG	CGCGCT	CGGCTACT	ACGAC	CGGCTG	CGGCTG	GGGCG	CGCCG	CAAGT	CCCCCG	1680
QY	1959	TACTGCG	GCACCA	AGTCCG	GGCTCT	GGTGT	GTGCTG	CTTGTG	CGCCCA	ACAGG	CGCGGCG	2018
Db	1681	TACTGCG	GCACCA	AGTCCG	GGCTCT	GGTGT	GTGCTG	CTTGTG	CGCCCA	ACAGG	CGCGGCG	1740

QY	2019	CGCGGCATATGGCGGGCGCCCAATTCCTTACTGTGGCGAGAGGCCCGAGGGGCTGTGGCCGGAG	2078
Db	1741	GGGGGCATATGGCGGGCGCCCAATTCCTTACTGTGGCGAGAGGCCCGAGGGGCTGTGGCCGGAG	1800
QY	2079	CGCTTCGCGCGCTGCGCGCGCGCGCGCGCGCGAGGAGCCCAAGGCCCAAGGACTGTCCGATTTC	2138
Db	1801	CGCTTCGCGCGCTGCGCGCGCGCGCGCGCGAGGAGCCCAAGGCCCAAGGACTGTCCGATTTC	1860
QY	2139	AGCTGGATCGAGACGCGCTCTCTCGATCAATGCATCTTCAGCGACTGTGGGGATTTC	2198
Db	1861	AGCTGGATCGAGACGCGCTCTCTCGATCAATGCATCTTCAGCGACTGTGGGGATTTC	1920
QY	2199	GAGCAGGCGCAAGCGGAGGCGGATTCGCGCGCGCGACAAGCGCGGTCTCGAGATTCGTC	2258
Db	1921	GAGCAGGCGCAAGCGGAGGCGGATTCGCGCGCGCGACAAGCGCGGTCTCGAGATTCGTC	1980
QY	2259	CGCGCTCAAGACGAGGAGTGCTGGGCCAGCGGAGCTGCAGAGAACTGCGCCCAAGGACATT	2318
Db	1981	CGCGCTCAAGACGAGGAGTGCTGGGCCAGCGGAGCTGCAGAGAACTGCGCCCAAGGACATT	2040
QY	2319	AGCGGCTATATATGGCTTCTACTCGGACAGCTAAGCCGCGCTGCGCGCGCGCGCGCGCG	2378
Db	2041	AGCGGCTATATATGGCTTCTACTCGGACAGCTAAGCCGCGCTTACCGCGCGCGCGCGCGCG	2100
QY	2379	CGGCGCGGACCGCCCGAGCGCGCGCTTCACAGCTCTTCCGACGCTCGGCTCCCGCACACTCC	2438
Db	2101	CGGCGCGGACCGCCCGAGCGCGCGCTTCACAGCTCTTCCCGACGCTCGGCTCCCGCACACTCC	2160
QY	2439	TCCTTTCGCGACCACTCATTTTATTGACCGCTCGATGGCGGTCTGCAGCGAATATGTCA	2498
Db	2161	TCCTTTCGCGACCACTCATTTTATTGACCGCTCGATGGCGGTCTGCAGCGAATATGTCA	2220
QY	2499	GGTCTCCGAGGGATTTTAACCTTTTGTGACAGCAGTCTTGCAATATAGCTCACCGAC	2558
Db	2221	GGTCTCCGAGGGATTTTAACCTTTTGTGACAGCAGTCTTGCAATATAGCTCACCGAC	2280
QY	2559	CTTCAACTTTGCTGTAAACCTTTTGATTTT	2588
Db	2281	CTTCAACTTTGCTGTAAACCTTTTGATTTT	2310
RESULT 3			
LOCUS	CQ769652	1793 bp	DNA
DEFINITION	Sequence 378 from Patent WO2003058021.	linear	PAT 04-MAR-2004
ACCESSION	CQ769652		
VERSION	CQ769652.1	GI:45114182	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE			
AUTHORS	1		
TITLE	Koenig-Hofman,K., Kazinski,M., Schaefer,R. and Kesper,B.		
JOURNAL	Novel apoptosis-inducing dna sequences		
	Patent: WO 2003058021-A 378 17-JUL-2003;		
	Xantos Biomedicine AG (DE)		
FEATURES			
source	Location/Qualifiers		
	1..1793		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	56.3%;	Score 1793;	DB 6; Length 1793;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1793;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	225	GTCTGTGATTTCTAGTTTATGATTAATAGACTTTTAAACCAAGGACGGGAGGCGAG	284
Db	1	GTCTGTGATTTCTAGTTTATGATTAATAGACTTTTAAACCAAGGACGGGAGGCGAG	60

QY 285 TGTTCAGGTTCTAGAGCTATGCACTGAGACCTGCTCTTCTCTTCTTATCATGCTCTCC 344
 DB 61 TGTTCAGGTTCTAGAGCTATGCACTGAGACCTGCTCTTCTCTTCTTATCATGCTCTCC 120
 QY 345 AAGAAATTTCTCAATGATGAGAGAGAGCTACCCACATTCAGGCGGATTCGAGCTTCTTG 404
 DB 121 AAGAAATTTCTCAATGATGAGAGAGAGCTACCCACATTCAGGCGGATTCGAGCTTCTTG 180
 QY 405 CACGATCATCCCATTTATCTGACCACTGACCACTGAGAGAGATTCACCTTTGAAAAA 464
 DB 181 CACGATCATCCCATTTATCTGACCACTGACCACTGAGAGAGATTCACCTTTGAAAAA 240
 QY 465 ATTACAGGCGGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 524
 DB 241 ATTACAGGCGGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 525 CCAAGGAGAGCTGACAGAGAGATTAACCTCTCTGATCTTGAAGGAGCTCTGAGCTTCTG 584
 DB 301 CCAAGGAGAGCTGACAGAGAGATTAACCTCTCTGATCTTGAAGGAGCTCTGAGCTTCTG 360
 QY 585 CACAGTTTCGATGAGCTCTGCTGAGATGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 644
 DB 361 CACAGTTTCGATGAGCTCTGCTGAGATGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
 QY 645 TCTGCGGAGCTGCT 704
 DB 421 TCTGCGGAGCTGCT 480
 QY 705 CCGGCTCTTCT 764
 DB 481 CCGGCTCTTCT 540
 QY 765 GAGAGCTTCAACAGAGCT 824
 DB 541 GAGAGCTTCAACAGAGCT 600
 QY 825 CCGTACCCACAGAGATGAGAGCT 884
 DB 601 CCGTACCCACAGAGATGAGAGCT 660
 QY 885 ACCGACCCGAGAGATGAGAGCT 944
 DB 661 ACCGACCCGAGAGATGAGAGCT 720
 QY 945 CTGAAATTTTCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004
 DB 721 CTGAAATTTTCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 1005 CTTTTTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1064
 DB 781 CTTTTTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 840
 QY 1065 GATGATGATTTTGGCGGATTCCTCAATCACTGAGAGGTTTCAAGAGGCAATGGGTTCTTGC 1124
 DB 841 GATGATGATTTTGGCGGATTCCTCAATCACTGAGAGGTTTCAAGAGGCAATGGGTTCTTGC 900
 QY 1125 GGCAGAGCGGACCAATGATGAGAGAGATTCGAGTCTATATGATCCGATTCCTCCCAAC 1184
 DB 901 GGCAGAGCGGACCAATGATGAGAGAGATTCGAGTCTATATGATCCGATTCCTCCCAAC 960
 QY 1185 ACTGAGGCTCACTGATGAGCGGCAAGAAATCTCTTTGAAAAATTTAAATTTTAAATTTT 1244
 DB 961 ACTGAGGCTCACTGATGAGCGGCAAGAAATCTCTTTGAAAAATTTAAATTTTAAATTTT 1020
 QY 1245 AAGAGAGCTTCAATTAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1304
 DB 1021 AAGAGAGCTTCAATTAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 QY 1305 CCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1364
 DB 1081 CCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1365 CGGCTGACAGAGCTTCACTTTCTCTGAGACTCATGTTCACTGCGCTGACCGGCTTACCAAC 1424

DB 1141 CGGCTGACAGAGCTTCACTTTCCCTGAGACTCAGTTCACTCCCGCTACCGGCTTACCAAGAC 1200
 QY 1425 AGGATATTACACACTGAAATATGATCACACCTTTTCAAAAGATTTGGGATTAAT 1484
 DB 1201 AGGATATTACACACTGAAATATGATCACACCTTTTCAAAAGATTTGGGATTAAT 1260
 QY 1485 TATGACAGATCTACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1544
 DB 1261 TATGACAGATCTACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 QY 1545 CCGGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1604
 DB 1321 CCGGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 1605 CAGTTCTGAGACACTACAGCCCAAGAGCCCGCTTCAACCGGCGGAGGCGGAGCCCGAG 1664
 DB 1381 CAGTTCTGAGACACTACAGCCCAAGAGCCCGCTTCAACCGGCGGAGGCGGAGCCCGAG 1440
 QY 1665 CCGGCTGACAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1724
 DB 1441 CCGGCTGACAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 1725 GACCGGAGCGGCGCTGCGCCGCAACGCTGTTTGTGACGCGGCGCAACACCGGCTGAGC 1784
 DB 1501 GACCGGAGCGGCGCTGCGCCGCAACGCTGTTTGTGACGCGGCGCAACACCGGCTGAGC 1560
 QY 1785 TTGCGGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1844
 DB 1561 TTGCGGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1845 TCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1904
 DB 1621 TCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
 QY 1905 CCGCTCGGCTACTACGCGGACCGGCTGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1964
 DB 1681 CCGCTCGGCTACTACGCGGACCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
 QY 1965 GGCACCAAGTGGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793
 DB 1741 GGCACCAAGTGGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793

RESULT 4
 BC029289
 LOCUS
 DEFINITION
 Homo sapiens T-box, brain, 1, mRNA (cdna clone IMAGE:5181773), partial cds.
 ACCESSION
 BC029289.1 GI:20379748
 VERSION
 KEYWORDS
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1793)
 Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Useth, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Viallon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Matra, M.A.
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

REFERENCE
 2 (bases 1 to 1793)
 NIH MGC Project
 Direct Submission
 Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-rt@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H., Kowals, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
 Series: IRAK Plate: 50 Row: C Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22547231.

FEATURES
 source
 Location/Qualifiers
 1..1793

organism="Homo sapiens"
 mol_type="mRNA"
 db_xref="taxon:9606"
 clone="IMAGE:5181773"
 tissue_type="Brain, Lung, Testis, adult, pooled whole"
 clone_lib="NIH MGC_115"
 lab_host="DH10B"
 note="Vector: pCMV-SPORT6"

1..51793
 gene
 /gene="TBR1"
 /note="synonym: TES-56"
 /db_xref="GeneID:10716"
 /db_xref="MIM:604616"
 79..>1793

codon_start=1
 /product="TBR1 protein"
 /protein_id="AAH29289.1"
 /db_xref="GI:62089470"
 /db_xref="GeneID:10716"
 /db_xref="MIM:604616"
 /translation="MOLEHCLEPSITMLSKPLNVSSYPHSGSELVLDHPITSTTDNLERSSPLKIRITRMNTDNTNPDNSKSDPDVORSKLSPLDGESELRHSFDSAA DRYLSQSSQPOSAATAPSAMFPYGOHPAHAFSISGSPRYMAHHVITNGAANSL LNSSPQGYPTGAPYPOQYSGOAPYQPSSTPGVLPKAOVYLGNRLMKPH RHOTEMILITKGRBMPPLSFNLSGLDIPAHYNITVDVLADPNHMRPGGKWWPCGK ADTNVQGNRYVYHPSPNTGAMWROEISFGKLKLTNNKGSANNNGQNVYLOSLLKYO PLHVAEVEDGIEDTSOPGRVQTFPFTQFLAVAYONTDIQLKIDNPFAGFR DNYVDITGCDMDRLTPSPNDSPRSQIVGARVYAGSFLDQFVSNYAKAFHFGAG ASPGPTDRSVHTNGLISPOAEDPGASPORWFTVTPANNLIDPAAAYATDTAFAG NATLISYAAGVYKALPLQAGACTGRPLGYADPSGWARSPQYCGTKSGSVLPCWP NSAAA"

ORIGIN

Query Match 56.3%; Score 1793; DB 8; Length 1793;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	225	GTCTGTGATTTCTAGTTTATGATAAATAGACTTTTAAAAACGAGGACGGAGGCGAG	284
DB	1	GTCTGTGATTTCTAGTTTATGATAAATAGACTTTTAAAAACGAGGACGGAGGCGAG	60
QY	285	TGTTCAAGTCTAGAGCTATGACGCTGAGACCTCTTCTCTTATCATGCTCTCC	344
DB	61	TGTTCAAGTCTAGAGCTATGACGCTGAGACCTCTTCTCTTATCATGCTCTCC	120
QY	345	AAGAATTTTCATGTGAGAGAGCTACCCACATTCAGGCGGATCCGAGCTTGTCTG	404
DB	121	AAGAATTTTCATGTGAGAGAGCTACCCACATTCAGGCGGATCCGAGCTTGTCTG	180
QY	405	CAGCATATCCCATTTATCTGACCACTGACCAACTGAGAGAGATTACCTTTGAAAAA	464
DB	181	CAGCATATCCCATTTATCTGACCACTGACCAACTGAGAGAGATTACCTTTGAAAAA	240
QY	465	ATTACAGGGGAGTACGAAATCAGTCAAGATACAGACAAATTTTCTGACTCCAGACTCA	524
DB	241	ATTACAGGGGAGTACGAAATCAGTCAAGATACAGACAAATTTTCTGACTCCAGACTCA	300
QY	525	CCAGGGGACGTCCAGAGAAATCTCTCTCTGTTGACGGGGTCTTGAGCTTCTG	584
DB	301	CCAGGGGACGTCCAGAGAAATCTCTCTCTGTTGACGGGGTCTTGAGCTTCTG	360
QY	585	CACAGTTTCAGTGGCTGCTGACAGATCGCTACCTCTCTCTGCTCAGTCCAGCCACAG	644
DB	361	CACAGTTTCAGTGGCTGCTGACAGATCGCTACCTCTCTCTGCTCAGTCCAGCCACAG	420
QY	645	TCTGCGGCACTGCTCCAGTGCATGTCCTGATCCCTGATCCCGGCGACGACCGGCGAC	704
DB	421	TCTGCGGCACTGCTCCAGTGCATGTCCTGATCCCGGCGACGACCGGCGCGAC	480
QY	705	CCCCCTTCTCATTCGGGACGCTTACGCGCTACATGCGCCACCCCGGTTCATACCAAC	764
DB	481	CCCCCTTCTCATTCGGGACGCTTACGCGCTACATGCGCCACCCCGGTTCATACCAAC	540
QY	765	GAGAGCTTACAGAGCTCTGTCCTGTCCTGCGGAGGATACCCCAAGCGGCGCTAC	824
DB	541	GAGAGCTTACAGAGCTCTGTCCTGTCCTGCGGAGGATACCCCAAGCGGCGCTAC	600
QY	825	CCCTACCCACAGCAGTACGGGCACTCTCCACCAAGAGCTCCGTTCTACCAAGTTCTCTCC	884
DB	601	CCCTACCCACAGCAGTACGGGCACTCTCCACCAAGAGCTCCGTTCTACCAAGTTCTCTCC	660
QY	885	ACCCAGCCGGGGCTGGTGGCCCGGCAAGACAGGTGTAACCTGTGCACAGGCCCTTTGG	944
DB	661	ACCCAGCCGGGGCTGGTGGCCCGGCAAGACAGGTGTAACCTGTGCACAGGCCCTTTGG	720
QY	945	CTGAATTTTCACCGGCAACCAAGGATGATCATACCAACAGGAAAGGCGCATGTT	1004
DB	721	CTGAATTTTCACCGGCAACCAAGGATGATCATACCAACAGGAAAGGCGCATGTT	780
QY	1005	CTTTTAAATTTTAAATTTCTGCTCGATCCCAAGCTCATTAACAATTTTTTGTG	1064
DB	781	CTTTTAAATTTTAAATTTCTGCTCGATCCCAAGCTCATTAACAATTTTTTGTG	840
QY	1065	GATGTGATTTTGGGAGATCCCAATCATGAGAGTTTCAAGAGCAATGGGTCTTGGC	1124
DB	841	GATGTGATTTTGGGAGATCCCAATCATGAGAGTTTCAAGAGCAATGGGTCTTGGC	900
QY	1125	GGCAAGGAGGACCAATGTGCAAGAAATCGGCTATATGATCCGATTTCCCAAC	1184
DB	901	GGCAAGGAGGACCAATGTGCAAGAAATCGGCTATATGATCCGATTTCCCAAC	960
QY	1185	ACTGGGGCTCACTGGATGCGCAAGAAATCTTTTGGAAAAATTAAACTTACGAACAC	1244
DB	961	ACTGGGGCTCACTGGATGCGCAAGAAATCTTTTGGAAAAATTAAACTTACGAACAC	1020
QY	1245	AAAGAGCTTCAATTAACATAGGAGAGTGTGTTTACAGTCTTGCAAGTACAG	1304
DB	1021	AAAGAGCTTCAATTAACATAGGAGAGTGTGTTTACAGTCTTGCAAGTACAG	1080

QY	1305	CCCCGCGTCAATGATGGTGAAGTAAAGAGGACGGACCGAGGACATATGACAGCCGGG	1364
Db	1081	CCCCGCGTCAATGATGGTGAAGTAAAGAGGACGGACCGAGGACATATGACAGCCGGG	1364
QY	1365	CGCGTGCAGAGCTTTCATCTTTCCTGAGACTGATTCTATTCGCGCTGACCGCTTACCAATC	1424
Db	1141	CGCGTGCAGAGCTTTCATCTTTCCTGAGACTGATTCTATTCGCGCTGACCGCTTACCAATC	1424
QY	1425	ACGGATATTACAACTGAAAAATGATTCACAACTCTTTTGCAAAAGATTTCCGGGATTAAT	1484
Db	1201	ACGGATATTACAACTGAAAAATGATTCACAACTCTTTTGCAAAAGATTTCCGGGATTAAT	1484
QY	1485	TATGACACAGATCTTACACCGGCTGTGACATGGAACCGCTGACACCCCTGCGCCCAAGACTCG	1544
Db	1261	TATGACACAGATCTTACACCGGCTGTGACATGGAACCGCTGACACCCCTGCGCCCAAGACTCG	1544
QY	1545	CCGGGCTGCAGATGTGTCGCCGGGGCCGGCTACGACATGGCGGCTCTTCTCTCAGAGAC	1604
Db	1321	CCGGGCTGCAGATGTGTCGCCGGGGCCGGCTACGACATGGCGGCTCTTCTCTCAGAGAC	1604
QY	1605	CAGTTTCGTGAGCAACTACGCCAAGGCCCGCTTTCACACCGGCGCGCGGCGCGGCGCGG	1664
Db	1381	CAGTTTCGTGAGCAACTACGCCAAGGCCCGCTTTCACACCGGCGCGCGGCGCGGCGCGG	1664
QY	1665	CCGGGATACGACCCGACGCTGTCGCGACACCAAAGGCTGTGTCGCGCGACAGCCGAG	1724
Db	1441	CCGGGATACGACCCGACGCTGTCGCGACACCAAAGGCTGTGTCGCGCGACAGCCGAG	1724
QY	1725	GACCCGGGCGGCGCCCTGCGCGCAACGCGTGTGTTGTGACGCGCGGCAACAAACCGGCTGGAC	1784
Db	1501	GACCCGGGCGGCGCCCTGCGCGCAACGCGTGTGTTGTGACGCGCGGCAACAAACCGGCTGGAC	1784
QY	1785	TTCCGCGGCTCGGCTTATGACACCGGCAACCGGACTTTCGCGGCGCAACGCGGACGCTGCTC	1844
Db	1561	TTCCGCGGCTCGGCTTATGACACCGGCAACCGGACTTTCGCGGCGCAACGCGGACGCTGCTC	1844
QY	1845	TCCTTACGCGCGCGCGGCGGTGAAGCGCTGCCGCTGCAGAGCTGCAGAGCTTGCATGTGCGCGC	1904
Db	1621	TCCTTACGCGCGCGCGGCGGTGAAGCGCTGCCGCTGCAGAGCTGCAGAGCTTGCATGTGCGCGC	1904
QY	1905	CCGCTCGGCTACTAAGCCGACCCCGTTCGCGGCTGGGGGCGCCGAGTCCCGCGAGTACGC	1964
Db	1681	CCGCTCGGCTACTAAGCCGACCCCGTTCGCGGCTGGGGGCGCCGAGTCCCGCGAGTACGC	1964
QY	1965	GGCACCAAGTCGGGCTCGGCTGTGCTCGCTGCTGGCCCAACAGCGCGCGCGCGCGC	2024
Db	1741	GGCACCAAGTCGGGCTCGGCTGTGCTCGCTGCTGGCCCAACAGCGCGCGCGCGCGCGC	2024

RESULT 5				
AC009487				
LOCUS	AC009487	183666 bp	DNA	linear
DEFINITION	Homo sapiens BAC clone RP11-334E15 from 2,			PRI 21-APR-2005
ACCSSION	AC009487			
VERSION	AC009487.3	GI:7321968		
KEYWORDS	HTG.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homnidae; Homo.			
REFERENCE	Hou,S., Maupin,R. and Scott,S.			
AUTHORS	Title The sequence of Homo sapiens BAC clone RP11-334E15			
JOURNAL	Unpublished (2001)			
REFERENCE	2 (bases 1 to 183666)			
AUTHORS	Waterston,R.H.			
TITLE	Direct Substistion			
JOURNAL	Submitted (24-AUG-1999) Genome Sequencing Center, Washington			
	University School of Medicine, 4444 Forest Park Parkway, St. Louis,			
	MO 63108, USA			
REFERENCE	3 (bases 1 to 183666)			
AUTHORS	Waterston,R.H.			

JOURNAL	TITLE	Direct Submission
	Submitted (24-MAR-2000)	Genome Sequencing Center, Washington
	University School of Medicine,	4444 Forest Park Parkway, St. Louis,
	MO 63108, USA	
REFERENCE	4 (bases 1 to 183666)	
AUTHORS	Waterston,R.H.	
TITLE	Direct Submission	
JOURNAL	Submitted (04-APR-2000)	Genome Sequencing Center, Washington
	University School of Medicine,	4444 Forest Park Parkway, St. Louis,
	MO 63108, USA	
REFERENCE	5 (bases 1 to 183666)	
AUTHORS	Waterston,R.H.	
TITLE	Direct Submission	
JOURNAL	Submitted (07-APR-2000)	Genome Sequencing Center, Washington
	University School of Medicine,	4444 Forest Park Parkway, St. Louis,
	MO 63108, USA	
REFERENCE	6 (bases 1 to 183666)	
AUTHORS	Waterston,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (30-SEP-2000)	Department of Genetics, Washington
	University, 4444 Forest Park Avenue, St. Louis,	Missouri 63108, USA
REFERENCE	7 (bases 1 to 183666)	
AUTHORS	Wilson,R.K.	
TITLE	Direct Submission	
JOURNAL	Submitted (21-APR-2005)	Genome Sequencing Center, Washington
	University School of Medicine,	4444 Forest Park Parkway, St. Louis,
	MO 63108, USA	
COMMENT	On Mar 24, 2000 this sequence version replaced gi:6042135.	

NOTICE

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

The RPT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Prengren, E., Tatematsu, M., Catenease, J. J. and de Jong, P. J. (1988) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Vries and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-26B22, 200 base pair overlap. Actual start of this clone is at base position 166089 of RP11-26B22; actual end of this clone is at base position 183666 of RP11-334E15.

```
FEATURES      location/Qualifiers
source        1. .18366
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
```

gene /chromosome="2" /clone="RP11-334E15" /clone_id="RPCT-11" 12580..38516 /gene="PSMD14" join(12580,12687,18220,18294,22194,22319,35953,36015,38418,38516) /gene="PSMD14" join(<12580,12687,18220,18294,22194,22319,35953,36015,38418,38516) /gene="PSMD14" /note="Homo sapiens proteasome (prosome, macropain) 26S subunit, non-ATPase, 14 (PSMD14), mRNA.; H_NH0334E15.1 This gene was based on gi(5031980) Continued from H_NH0026B22.2" /codon_start=1 /product="unknown" /protein_id="AAV15016.1" /db_xref="gi:52822468" /translation="VVIDAFRLINANNMVLGHERPQTTSNLGHLNKSIOALLHGLNR HYSVITLYNRKNELEQKMLINLHKRSWMEGLTLDYSEHCNHSVVKEMLELAKNN KAVSEEDMPTEQEAIKNVGQKQDPKRHLSEHVLIAMTSNIYQCAALMLDVFK"

misc_feature 41494..42018 /note="CpG_island (%GC=73.5, o/e=0.71, #CpGs=58)" 43225..52178 /gene="TBRI" join(43225,44218,44792,44946,45317,45438,46008,46166,47312,47373,50485,52178) /gene="TBRI" join(43527,44218,44792,44946,45317,45438,46008,46166,47312,47373,50485,51343) /gene="TBRI" /note="Homo sapiens T-box, brain, 1 (TBRI), mRNA.; H_NH0334E15.2 This gene was based on gi(22547231)" /codon_start=1 /product="unknown" /protein_id="AAV15017.1" /db_xref="gi:52822469" /translation="WQLEHDCSPSILMSKFLANVSSSYPHSGSELVLDHPILISTTD NLRESPLKTIIRGTNGSDTDFNPDSKDSGDYQSRKSLVLDGSELRSFPGSAA DRYLISOSSQPSASATAPASAMPYRQGGPHAPAFSGSPSRVMAHHPIVINAYNL LSNSPQGYPTAGYPTPOQYGHSTQGLAFYPTSTQPELVGKAQVYLCNRPMLKXH RHQEMITIKQGRMPPLPSFNISGLDEPTAHYNI.FVDTLADPNHMFQGGKWPCK ADITVQGNRYVMHPDSPTGTAMMMROEISFEKLLTNNGKASNNNGQMVVLOSIAHKQ PRLHVNEVNEDEGTEDSQPGAVMOTFEPEFTIAYTAAYMTDILQIKDHNPEAKGR DNVPITLYSCMDRLTLPSPDSRPSQIYPGARVMAQSFLDDQCVSNVAKRFPVGG AGRPGRDYSVPMHTNGLSLSPQADRPAPRSQKRFVYPPANRRDLPAASADTAPDPA NATUILLSTAAAGVAKALPLQAGCTGRPLGYIADRSWGARSPPDYCGTSGSVALPCP NSAAAAAMAGANPYLGEAEAGLAERSLPLPGAEEAPKPKDLSDSSWIEFTPSIKSI DSSAGTIEQAKRRRISPADTPVSESSPLTSEVTLAQRDCEKCAKDISGYGFSHS"

misc_feature 43900..44330 /gene="TBRI" /note="CpG_island (%GC=63.3, o/e=0.70, #CpGs=30)" 45767..46201 /gene="TBRI" /note="CpG_island (%GC=67.8, o/e=0.83, #CpGs=45)" 50441..51502 /gene="TBRI" /note="CpG_island (%GC=71.4, o/e=0.85, #CpGs=134)" 54187..55340 /note="CpG_island (%GC=62.6, o/e=0.81, #CpGs=99)" 91821..92098 /note="CpG_island (%GC=60.1, o/e=0.83, #CpGs=20)" 134790..135005 /note="CpG_island (%GC=77.8, o/e=0.97, #CpGs=37)" 172044..172317 /note="CpG_island (%GC=60.6, o/e=0.78, #CpGs=21)"

ORIGIN

Query Match 53.2%; Score 1694; DB 8; Length 183666; Best Local Similarity 100.0%; Pred. No. 0;

Matches 1694: Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
OY	1493	GATCTACACCGGCTGTGACATGGA	CGCGCTGACCCCTCGCCCAACGACTCGCGGCTC	1552					
Db	50485	GATCTACACCGGCTGTGACATGGA	CCCGCTGACCCCTCGCCCAACGACTCGCGGCTC	5054					
OY	1553	GCACATGTGTCGCGGGGCGCGGCTACG	CGCATGGCGCGCTCTTCTGTGAGAACGAGTTGCT	1612					
Db	50545	GCACATGTGTCGCGGGGCGCGGCTACG	CGCATGGCGCGCTCTTCTGTGAGAACGAGTTGCT	5060					
OY	1613	GAGCACTACGCGCAAGGCGCGCTTCC	ACCGGGCGCGGGCGCGGCGCCCGGGCGGGTAC	1672					
Db	50605	GAGCACTACGCGCAAGGCGCGCTTCC	ACCGGGGGGGGGCGCGGCGCCCGGGCGGGTAC	5066					
OY	1673	GAACCGCAGCGTGCACACACCA	CGGGCTGTCTGCGCGACGAGCCGACCTGGG	1732					
Db	50665	GAACCGCAGCGTGCACACCA	CGGGCTGTCTGCGCGACGAGCCGACCTGGG	5072					
OY	1733	CGGCGCTTCGCGCGAAGCTGTGTTTG	AGCGCGGCGCAACACCGGTGACCTTCGGGCG	1792					
Db	50725	CGGCGCTTCGCGCGAAGCTGTGTTTG	AGCGCGGCGCAACACCGGTGACCTTCGGGCG	5078					
OY	1793	CTCGGCGCTATGACACGCGCAC	AGCACTTCGCGGGCAACGCGACGCTGCTTTACG	1852					
Db	50785	CTCGGCGCTATGACACGCGCAC	AGCACTTCGCGGGCAACGCGGCAACGCTGCTTTACG	5084					
OY	1853	GGCGGCGGGCGTGAAGCGCTGCCG	CTGACAGCTGACGAGCTGACCTGAGCCGCGCTCG	1912					
Db	50845	GGCGGCGGGCGTGAAGCGCTGCCG	CTGACAGCTGACGAGCTGACCTGAGCCGCGCTCG	5090					
OY	1913	CTACTACGCGGACCCGTCGGGCTGT	GGGCGCCCGACATCTGCGGACCA	1972					
Db	50905	CTACTACGCGGACCCGTCGGGCTGT	GGGCGCCCGACATCTGCGGACCA	5096					
OY	1973	GTGCGGCTCGGTGTGCTCTGCTGAG	CCCAACAGCGCGCGGCGCGCGGACATGCGCG	2032					
Db	50965	GTGCGGCTCGGTGTGCTCTGCTGAG	CCCAACAGCGCGCGGCGCGCGGACATGCGCG	5102					
OY	2033	CGCCAACTCCCTACCTGCGGCGAG	AGGCGCTGGCGCGGACGCTCGCCGCTGCC	2092					
Db	51025	CGCCAACTCCCTACCTGCGGCGAG	AGGCGCTGGCGCGGACGCTCGCCGCTGCC	5108					
OY	2093	GCCCGGCGCGCGCGAGACCGCA	AGCCCAAGACTCTGTCCGATTCGAGCTGAGTCCG	2152					
Db	51085	GCCCGGCGCGCGCGAGACCGCA	AGCCCAAGACTCTGTCCGATTCGAGCTGAGTCCG	5114					
OY	2153	GCCTTCCTCGATCAAGTGCATTCGA	CTCCAGGACTCGGGATTTCAGACGCGCAAGCG	2212					
Db	51145	GCCTTCCTCGATCAAGTGCATTCGA	CTCCAGGACTCGGGATTTCAGACGCGCAAGCG	5120					
OY	2213	GAGCGGATCTCGCGCGCGCA	CGCCCGTGTCCGAGAGTTTCGCCCGCTCAAGACGA	2272					
Db	51205	GAGCGGATCTCGCGCGCGCA	CGCCCGTGTCCGAGAGTTTCGCCCGCTCAAGACGA	5126					
OY	2273	GGTCTGCGCCACGCGGACTGCGAG	AAAGAACTGCGCCAGACATTAGCGGCTATATG	2332					
Db	51265	GGTCTGCGCCACGCGGACTGCGAG	AAAGAACTGCGCCAGACATTAGCGGCTATATG	5132					
OY	2333	CTTCTACTCGACAGCTAAGGCGCG	CGCGCTGCGCGCGCGCGCGCGCGCGCGGACCCCG	2392					
Db	51325	CTTCTACTCGACAGCTAAGGCGCG	CGCGCTGCGCGCGCGCGCGCGCGCGCGGACCCCG	5138					
OY	2393	AGCAGCGCCCTCACAGCTCTTCC	CCAGCTCGCCTCCCAACACTCTCTTGTGGCACCA	2452					
Db	51385	AGCAGCGCCCTCACAGCTCTTCC	CCAGCTCGCCTCCCAACACTCTCTTGTGGCACCA	5144					
OY	2453	CTCATTTTATTTGA	CCCTCGATGGCGGTCTGACGCAATTAAGTACGAGTCTTCGAGCGTG	2512					
Db	51445	CTCATTTTATTTGA	CCCTCGATGGCGGTCTGACGCAATTAAGTACGAGTCTTCGAGCGTG	5150					
OY	2513	ATTTTAACTTTTGTGACAGCA	AGTCTCTGCAATTAAGTACACGCACTTTCGACTG	2572					
Db	51505	ATTTTAACTTTTGTGACAGCA	AGTCTCTGCAATTAAGTACACGCACTTTCGACTG	5156					

[illegible]

RESULT 6
AC165777/c

LOCUS	AC165777	101040 bp	DNA	linear	HTG 17-JUL-2005
DEFINITION	Bos taurus clone CH240-161a23, *** SEQUENCING IN PROGRESS ***, 27				
ACCESSION	AC165777				
VERSION	AC165777.1	GI:70912797			
KEYWORDS	HTG; HTGS_PHASE1.				
SOURCE	Bos taurus (cow)				
ORGANISM	Bos taurus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;				
	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;				
	Pecora; Bovidae; Bovinae; Bos.				
	1 (bases 1 to 101040)				
REFERENCE	Muzny, D., Adams, C., Agbdi II, O., Allen, C., Albrooke, S., Archer, P.,				
AUTHORS	Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltzan, R.,				
	Berducci, A., Biswal, K., Blyth, P., Bonham, H., Bunyah, C., Butch, P.,				
	Cadoore, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,				
	Chakoe, J., Chahrouh, M., Chavez, D., Chen, A., Chen, G., Chen, R.,				
	Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,				
	Curry, S., Dai, W., Davila, M. L., Davis, C., Day-Carroll, L., De				
	Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, H.,				
	Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durkin, C.,				
	Dzinda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,				
	Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,				
	Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,				
	Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,				

Gonzalez-Garay, M., Guevarra, W., Holder, M.B., Haaland, W., Heeberlein, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L.T., Huylk, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafatis, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kwis, A., Kwis, C., Lara, F., Leal, S., Lee, K., Lee, S., Legall, F.I., Lemon, S., Lewis, L., Li, B., Lin, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyang, D., London, P., Lopez, J., Lorenshuwa, L., Lozada, R., Luk, T., Madu, R., Maheswari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McCelland, H., McPherson, J., Mercado, C., Metzker, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Mundasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastang, E., Nott, A., Nwokediemeh, O., Obregon, M., Ochi-Okorie, C., Oden, E., Okunonu, G., Okunonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Pitman, E., Pui, L., Puzo, M., Qin, X., Quinn, A., Quiroz, J., Rabats, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W., Santibanez, V., Santos, R., Saverly, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatok, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Umani, K., Varco, C., Verduzco, D., Villanasa, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.

Direct Submission
 2 (bases 1 to 101040)

Worley, K.C.
 Direct Submission
 Submitted (17-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center -----
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help.tmc.edu

----- Project Information -----
 Center project name: RJUJ
 Center clone name: CH240-161A23

----- Summary Statistics -----
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 98251 bases at least Q40
 Consensus quality: 103096 bases at least Q20
 Consensus quality: 106131 bases at least Q20
 Estimated insert size: 124067; sum-of-contigs estimation
 Estimated insert size: 107277; agarose-fp estimation
 Quality coverage: 2x in Q20 bases; agarose-fp estimation
 Quality coverage: 2x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length -----
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drafter_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 27 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	4966:	contig of 4966 bp in length
*	4967	5066: gap of unknown length
*	5067	7581: contig of 2515 bp in length
*	7582	7681: gap of unknown length
*	7682	11922: contig of 4241 bp in length
*	11923	12022: gap of unknown length
*	12023	16702: contig of 4680 bp in length

16703	16802: gap of unknown length	16703	16802: gap of unknown length
16803	16881: contig of 2079 bp in length	16803	16881: contig of 2079 bp in length
16882	18981: gap of unknown length	16882	18981: gap of unknown length
18982	22999: contig of 4018 bp in length	18982	22999: contig of 4018 bp in length
23000	25642: gap of unknown length	23000	25642: gap of unknown length
23100	25642: contig of 2543 bp in length	23100	25642: contig of 2543 bp in length
25643	25742: gap of unknown length	25643	25742: gap of unknown length
25743	31586: contig of 5844 bp in length	25743	31586: contig of 5844 bp in length
31587	31686: gap of unknown length	31587	31686: gap of unknown length
31687	34315: contig of 2629 bp in length	31687	34315: contig of 2629 bp in length
34316	34415: gap of unknown length	34316	34415: gap of unknown length
34416	37369: contig of 2954 bp in length	34416	37369: contig of 2954 bp in length
37370	37469: gap of unknown length	37370	37469: gap of unknown length
37470	39730: contig of 2261 bp in length	37470	39730: contig of 2261 bp in length
39731	39830: gap of unknown length	39731	39830: gap of unknown length
39831	42816: contig of 2986 bp in length	39831	42816: contig of 2986 bp in length
42817	42916: gap of unknown length	42817	42916: gap of unknown length
42917	48470: contig of 5554 bp in length	42917	48470: contig of 5554 bp in length
48471	48570: gap of unknown length	48471	48570: gap of unknown length
48571	51440: contig of 2870 bp in length	48571	51440: contig of 2870 bp in length
51441	51540: gap of unknown length	51441	51540: gap of unknown length
51541	53825: contig of 2285 bp in length	51541	53825: contig of 2285 bp in length
53826	53925: gap of unknown length	53826	53925: gap of unknown length
53926	56237: contig of 2312 bp in length	53926	56237: contig of 2312 bp in length
56238	56337: gap of unknown length	56238	56337: gap of unknown length
56338	59002: contig of 2665 bp in length	56338	59002: contig of 2665 bp in length
59003	59102: gap of unknown length	59003	59102: gap of unknown length
59103	67716: contig of 8614 bp in length	59103	67716: contig of 8614 bp in length
67717	67816: gap of unknown length	67717	67816: gap of unknown length
67817	71710: contig of 3894 bp in length	67817	71710: contig of 3894 bp in length
71711	71810: gap of unknown length	71711	71810: gap of unknown length
71811	74047: contig of 2237 bp in length	71811	74047: contig of 2237 bp in length
74048	74147: gap of unknown length	74048	74147: gap of unknown length
74148	77776: contig of 3629 bp in length	74148	77776: contig of 3629 bp in length
77777	77876: gap of unknown length	77777	77876: gap of unknown length
77877	86904: contig of 9028 bp in length	77877	86904: contig of 9028 bp in length
86905	87004: gap of unknown length	86905	87004: gap of unknown length
87005	89555: contig of 2551 bp in length	87005	89555: contig of 2551 bp in length
89556	89655: gap of unknown length	89556	89655: gap of unknown length
91999	91999: contig of 2344 bp in length	91999	91999: contig of 2344 bp in length
92000	92099: gap of unknown length	92000	92099: gap of unknown length
95461	95461: contig of 3362 bp in length	95461	95461: contig of 3362 bp in length
95462	95561: gap of unknown length	95462	95561: gap of unknown length
95562	97821: contig of 2260 bp in length	95562	97821: contig of 2260 bp in length
97822	97921: gap of unknown length	97822	97921: gap of unknown length
97922	101040: contig of 3119 bp in length.	97922	101040: contig of 3119 bp in length.
FEATURES	Location/Qualifiers	FEATURES	Location/Qualifiers
source	1..101040	source	1..101040
	/organism="Bos taurus"		/organism="Bos taurus"
	/mol_type="genomic DNA"		/mol_type="genomic DNA"
	/db_xref="taxon:9913"		/db_xref="taxon:9913"
	/clone="CH240-161A23"		/clone="CH240-161A23"
gap	4967..5066	gap	4967..5066
	/estimated_length=unknown		/estimated_length=unknown
gap	7582..7681	gap	7582..7681
	/estimated_length=unknown		/estimated_length=unknown
gap	1193..12022	gap	1193..12022
	/estimated_length=unknown		/estimated_length=unknown
gap	16703..16802	gap	16703..16802
	/estimated_length=unknown		/estimated_length=unknown
gap	18882..18981	gap	18882..18981
	/estimated_length=unknown		/estimated_length=unknown
gap	23000..23099	gap	23000..23099
	/estimated_length=unknown		/estimated_length=unknown
gap	25643..25742	gap	25643..25742
	/estimated_length=unknown		/estimated_length=unknown
gap	31587..31686	gap	31587..31686
	/estimated_length=unknown		/estimated_length=unknown
gap	34316..34415	gap	34316..34415
	/estimated_length=unknown		/estimated_length=unknown
gap	37370..37469	gap	37370..37469
	/estimated_length=unknown		/estimated_length=unknown
gap	39731..39830	gap	39731..39830
	/estimated_length=unknown		/estimated_length=unknown
gap	42817..42916	gap	42817..42916
	/estimated_length=unknown		/estimated_length=unknown
gap	48471..48570	gap	48471..48570
	/estimated_length=unknown		/estimated_length=unknown
gap	51441..51540	gap	51441..5154

QY 651 GCCACTGCTCCGAGTGGCATGTTCCCGTACCCCGGCGACGACGACCGG 700
Db 64776 GCCACTGCTCCGAGTGGCATGTTCCCGTACCCCGGCGACGACGACCGG 64727

RESULT 7
AR655897/c AR655897 514 bp DNA linear PAT 13-JUN-2005
LOCUS Sequence 37 from patent US 6893820.
DEFINITION AR655897
ACCESSION AR655897
VERSION AR655897.1 GI:67588385
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
TITLES 1 (bases 1 to 514)
Plasmid, C.
Detection of methylated CpG rich sequences diagnostic for malignant cells

JOURNAL
Patent: US 6893820-A 37 17-MAY-2005;
The Ohio State University Research Foundation; Columbus, OH

FEATURES
source
1..514
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 16.1%; Score 514; DB 6; Length 514;
Best Local Similarity 100.0%; Pred. No. 6.3e-288;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GCTGTGACATGACCGCGCTGACCCCTCGCCCAACGACTCGCGCGCTCGACATGTGC 1563
Db 514 GCTGTGACATGACCGCGCTGACCCCTCGCCCAACGACTCGCGCGCTCGACATGTGC 455
QY 1564 CCGGGGCGCGCTACGCGCATGCGCGCTCTTTCTTGACAGACCGATTCTGAGAACTTAG 1623
Db 454 CCGGGGCGCGCTACGCGCATGCGCGCTCTTTCTTGACAGACCGATTCTGAGAACTTAG 395
QY 1624 CCAAGGCGCGCTTCCACCGCGGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 1683
Db 394 CCAAGGCGCGCTTCCACCGCGGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 335
QY 1684 TGCAGACACCAACGAGGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1743
Db 334 TGCAGACACCAACGAGGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 275
QY 1744 CGCAACGCTGTTTGTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1803
Db 274 CGCAACGCTGTTTGTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 215
QY 1804 ACAACGCGCAACGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1863
Db 214 ACAACGCGCAACGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 155
QY 1864 TGAAGGCGCTGCGGCTGCAAGGCTGCACTGGCGCGCGCGCTGCTACGCGCG 1923
Db 154 TGAAGGCGCTGCGGCTGCAAGGCTGCACTGGCGCGCGCGCTGCTACGCGCG 95
QY 1924 ACCGTCGAGGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1983
Db 94 ACCGTCGAGGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 35
QY 1984 TGTGCGCGCTGCG 2017
Db 34 TGTGCGCGCTGCG 1

RESULT 8
BD024632 388 bp DNA linear PAT 27-AUG-2002
LOCUS BD024632
DEFINITION Sequence tag and encoded human protein.

ACCESSION BD024632
VERSION BD024632.1 GI:22565855
KEYWORDS
SOURCE JP 2001269182-A/878.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 388)
TITLES Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
JOURNAL Sequence tag and encoded human protein
Patent: JP 2001269182-A 878 02-OCT-2001;
GENSET

COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/878
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68 //G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS 117..386.

FEATURES
source
1..388
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 12.1%; Score 384; DB 6; Length 388;
Best Local Similarity 100.0%; Pred. No. 7.3e-212;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 TTGCTGTTGAAGTCTTTCTGTCTAGTGAAGGCGCTGTGATTTCTAGTTATGATTA 250
Db 5 TTGCTGTTGAAGTCTTTCTGTCTAGTGAAGGCGCTGTGATTTCTAGTTATGATTA 64
QY 251 ATAGGACTTTAAACACGAGGCGGAGGCGAGTGTCAAGTCTTAGAGCTATGACGT 310
Db 65 ATAGGACTTTAAACACGAGGCGGAGGCGAGTGTCAAGTCTTAGAGCTATGACGT 124
QY 311 GGAGCAGTCCCTTTCTCTTATCATGCTCTCCAAAGAAATTTCAATGTGACAGAG 370
Db 125 GGAGCAGTCCCTTTCTCTTATCATGCTCTCCAAAGAAATTTCAATGTGACAGAG 184
QY 371 CTACCCACATTGAGCGGATCCGAGCTTGTGACGATCATCCCATTAATCTGACAC 430
Db 185 CTACCCACATTGAGCGGATCCGAGCTTGTGACGATCATCCCATTAATCTGACAC 244
QY 431 TGACAACTGAGAGAGTTCACCTTGAAGAAATTTACAGGGGATGACGAATCAGTC 490
Db 245 TGACAACTGAGAGAGTTCACCTTGAAGAAATTTACAGGGGATGACGAATCAGTC 304
QY 491 AGATACAGCAATTTTCTGACTCCAGGACTCACAGGGAGCGTCCAGAGAACTAACT 550
Db 305 AGATACAGCAATTTTCTGACTCCAGGACTCACAGGGAGCGTCCAGAGAACTAACT 364
QY 551 CTCTCTGTTTGAAGCGGCTTC 574
Db 365 CTCTCTGTTTGAAGCGGCTTC 388

RESULT 9
AX885022 388 bp DNA linear PAT 18-DEC-2003
LOCUS AX885022
DEFINITION Sequence 885 from Patent EP1033401.
ACCESSION AX885022
VERSION AX885022.1 GI:40040767

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Dumas Mline Edwards J.B., Duclert A. and Giordano J.Y.
AUTHORS Expressed sequence tags and encoded human proteins
TITLE Patent: EP 1033401-A 885 06-SEP-2000;
JOURNAL Genset (FR)
FEATURES
source location/Qualifiers
1..388
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
117..>386
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE99136.1"
/db_xref="GI:40040768"
/translation="MQEHCLSPSITMSKFLNVSSSPHSGSELVHDPITITD
NLRSSPLKKITRGMNOSDITNFPDSKDSPPDVRSKLSPLVDGV"
ORIGIN
Query Match 12.1%; Score 384; DB 6; Length 388;
Best Local Similarity 100.0%; Pred. No. 7.3e-212; Indels 0; Gaps 0;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 TTGCTGGTTGAAGTCTTTCTGTCTAGTAGAGGGGGTCTGTGATTTCTAGTTTATGATTA 250
DB 5 TTGCTGGTTGAAGTCTTTCTGTCTAGTAGAGGGGGTCTGTGATTTCTAGTTTATGATTA 64
QY 251 ATTAGACTTTAAAAACAGAGGAGCGGAGCGGAGTTCAGGTTCTAGAGCTATGCACT 310
DB 65 ATAGAGCTTTAAAAACAGAGGAGCGGAGCGGAGTTCAGGTTCTAGAGCTATGCACT 124
QY 311 GGAGCACTGCTTTCTCTTCTATGCTCTCAAGAAATTTCTCAATGTAGAGCAG 370
DB 125 GGAGCACTGCTTTCTCTTCTATGCTCTCAAGAAATTTCTCAATGTAGAGCAG 184
QY 371 CTACCCACATTCAGGCGGATCCGAGCTTCTTTCAGATCATCCCATTAATCTGCAC 430
DB 185 CTACCCACATTCAGGCGGATCCGAGCTTCTTTCAGATCATCCCATTAATCTGCAC 244
QY 431 TGACAACTTGAGAGAGTTCACCTTTGAAAAAATTACAGGGGATGACGATCAGTC 490
DB 245 TGACAACTTGAGAGAGTTCACCTTTGAAAAAATTACAGGGGATGACGATCAGTC 304
QY 491 AGATACAGCAATTTTCTGACTCCAAAGACTCACAGGGGAGCGTCCAGAGAATTAAC 550
DB 305 AGATACAGCAATTTTCTGACTCCAAAGACTCACAGGGGAGCGTCCAGAGAATTAAC 364
QY 551 CTCTCCTGTCTTGACGGGGTCTC 574
DB 365 CTCTCCTGTCTTGACGGGGTCTC 388
RESULT 10
HSA338878/c 641 bp DNA linear PRI 18-JUL-2002
LOCUS HSA338878
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
ACCESSION AJ338878
VERSION AJ338878.1 GI:15883296
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 641)
Kutseiko A.S., Gizatullin R.Z., Al-Amin A.N., Wang F., Kvasha S.M.,

Podowski R.M., Matushkin Y.G., Gyanchandani A., Muravenko O.V.,
Levitsky V.G., Kolchanov N.A., Protopopov A.I., Kashuba V.I.,
Kiselev L.L., Wasserman W., Mahlestedt C. and Zabarovsky E.R.
NotI flanking sequences: a tool for gene discovery and verification
of the human genome
Nucleic Acids Res. 30 (14), 3163-3170 (2002)
JOURNAL
PUBMED 12136098
REFERENCE 2 (bases 1 to 641)
AUTHORS Zabarovsky E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumoriobiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
source location/Qualifiers
1..641
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NL6-GN7C"
ORIGIN
Query Match 11.9%; Score 380; DB 8; Length 641;
Best Local Similarity 100.0%; Pred. No. 1.8e-209; Indels 0; Gaps 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1638 CACCCGGGCGGGCGGGCGGGCGGGGCTACGAGCCGACGCGACACCAAC 1697
DB 380 CACCCGGGCGGGCGGGCGGGCGGGGCTACGAGCCGACGCGACACCAAC 321
QY 1698 GGGCTGTGTGCGCGCAGACGAGCCGAGACCAGGCGCGCGCTTCCGCGCAACGCTGTTT 1757
DB 320 GGGCTGTGTGCGCGCAGACGAGCCGAGACCAGGCGCGCGCTTCCGCGCAACGCTGTTT 261
QY 1758 GTGACGCCGGCAACAACCGGCTGAGCTTGGCGGCTCGGCTATGACAGGCCAGGAC 1817
DB 260 GTGACGCCGGCAACAACCGGCTGAGCTTGGCGGCTCGGCTATGACAGGCCAGGAC 201
QY 1818 TTCCGGGGCAACGGGCGCAGCCTGCTCTTACCGCGGCGGGCGTGAAGCCGCTCCG 1877
DB 200 TTCCGGGGCAACGGGCGCAGCCTGCTCTTACCGCGGCGGGCGTGAAGCCGCTCCG 141
QY 1878 CTGACAGCTGAGGCTGACCTGAGCCGCCCGCTGCTACTAGCCGACCCGTCGGCTGG 1937
DB 140 CTGACAGCTGAGGCTGACCTGAGCCGCCCGCTGCTACTAGCCGACCCGTCGGCTGG 81
QY 1938 GGGCGCCGCACTCCCCCGCACTTGGCGGCAAGTCGGGCTGGTGGCTGCTGG 1997
DB 80 GGGCGCCGCACTCCCCCGCACTTGGCGGCAAGTCGGGCTGGTGGCTGCTGG 21
QY 1998 CCAACAGCGCGCGCGCGC 2017
DB 20 CCAACAGCGCGCGCGCGC 1
RESULT 11
HSA324262/c 689 bp DNA linear PRI 18-JUL-2002
LOCUS HSA324262
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
ACCESSION AJ324262
VERSION AJ324262.1 GI:15868641
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 689)
Kutseiko A.S., Gizatullin R.Z., Al-Amin A.N., Wang F., Kvasha S.M.,
Podowski R.M., Matushkin Y.G., Gyanchandani A., Muravenko O.V.,
Levitsky V.G., Kolchanov N.A., Protopopov A.I., Kashuba V.I.,
Kiselev L.L., Wasserman W., Mahlestedt C. and Zabarovsky E.R.
NotI flanking sequences: a tool for gene discovery and verification

Db 181 CCTCAATCGTTGCATGAAATATTACTATGTGCCCTAATGCACACAATAAGCTA 235

Search completed: December 20, 2005, 21:44:11
Job time : 15686 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 12:35:58 ; Search time 1805 Seconds

(without alignment(s))
11763.829 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186
Sequence: 1 caggtgattatcctaataa.....caccacaacaccttaaaagg 3186

(Scoring table: OLIGO_NUC
Gapop_60.0, Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size: 0
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

N_Geneseq_21.*
1: geneseqn1908s:*
2: geneseqn1908s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	3186	100.0	3186	12	ADL97817 Human TBR
2	1793	56.3	1793	11	ABE179130 Human apo
3	1694	53.2	6878	5	ABA14582 Human ner
4	864	27.1	864	12	ACH87502 Human gen
5	515	16.2	515	12	ACH73791 Human gen
6	514	16.1	514	6	ABT07220 Human Cpg
7	384	12.1	388	3	AAC00887 Human sec
8	303	9.5	457	9	ACH15279 Human adu
9	281	8.8	463	13	ACF88226 Human STR
10	260	8.2	389	2	AAO39695 Expressed
11	260	8.2	389	2	AAO59107 Human bra
12	235	7.4	235	3	AAO18168 Human sec
13	60	1.9	60	6	ABN33794 Human spl
14	32	1.0	1026	6	ABQ22410 Oligonuc1
15	32	1.0	1026	6	ABQ22411 Oligonuc1
16	27	0.8	60	3	AAC19701 Human sec
17	26	0.8	2435	4	AAH31915 Human olf
18	26	0.8	2435	3	AAC79868 Human sec
19	26	0.8	6140	4	AAI05394 Human rep

C	20	26	0.8	6140	4	ABL98256	Ab198256 Human tes
C	21	26	0.8	6554	5	ABAI7685	Abai7685 Human ner
C	22	26	0.8	8691	4	ABL19838	Ab119838 Drosophi1
C	23	26	0.8	8746	4	ABL02984	Ab102984 Drosophi1
C	24	26	0.8	12237	6	ABL34359	Ab134359 Human imm
C	25	26	0.8	13321	4	AAS46422	Aas46422 Tumour su
C	26	26	0.8	15576	10	ADC86186	Adc86186 Human GPC
C	27	26	0.8	18376	4	AAK67167	Aak67167 Human imm
C	28	26	0.8	23249	6	AAS19608	Aas19608 Reference
C	29	26	0.8	23885	4	AAK73173	Aak73173 Human imm
C	30	26	0.8	23885	4	AAK70103	Aak70103 Human imm
C	31	26	0.8	25074	13	ABD32941	Abd32941 Human can
C	32	26	0.8	29836	12	ADO07012	Ado07012 NIMA-rela
C	33	26	0.8	32249	5	ABAI7155	Abai7155 Human ner
C	34	26	0.8	38575	11	ACN37227	Acn37227 Human per
C	35	26	0.8	57243	11	ACN44826	Acn44826 Human gen
C	36	26	0.8	90120	8	ADQ97413	Adq97413 Human can
C	37	26	0.8	110000	13	ABD32912_2	Continuation (3 of
C	38	26	0.8	110000	13	ABD32912_2	Continuation (3 of
C	39	26	0.8	138941	8	ACC79695	Acc79695 Human tum
C	40	26	0.8	143412	11	ACN44512	Acn44512 Mouse gen
C	41	26	0.8	143412	8	ABE280229	Ab280229 Human tra
C	42	26	0.8	249999	12	ADO20017	Adq20017 Human sof
C	43	26	0.8	260160	13	ADS99460	Ads99460 Murine ki
C	44	26	0.8	263852	13	ADP75180	Adp75180 Human End
C	45	26	0.8	304905	11		

ALIGNMENTS

RESULT 1	ADL97817	standard; DNA, 3186 BP.
XX	ID	ADL97817;
XX	ADL97817;	(first entry)
DT	03-JUN-2004	
XX	Human TBR1 DNA	SeqID 3.
DE	human; gene; ds; neuropsychiatric disorder; bipolar disorder;	
XX	schizophrenia; alpha-type II calcium/calmodulin dependent protein kinase;	
KW	CMKMT1-alpha; TBR1; Brachyury; neuroleptic; antidepressant; gene therapy;	
KW	mental disorder; mood disorder; psychosis; major depression.	
XX	Homo sapiens.	
OS	Homo sapiens.	
FX	Key	Location/Qualifiers
FT	CDS	303..2351
FT		/*tag= a
FT		/product= "TBR1 protein"
XX	MO2004020455-A2.	
XX	11-MAR-2004.	
XX	26-AUG-2003; 2003WO-US026873.	
PF	28-AUG-2002; 2002US-0406879P.	
PR	27-FEB-2003; 2003US-0451306P.	
XX	(STRD) UNIV IELAND STANFORD JUNIOR.	
PA	Burney WR, Jones EG, Molnar M;	
PI	WPI: 2004-239152/22.	
DR	P-PSDB; ADL97818.	
XX	Determining whether a subject has or is predisposed to a mental disorder,	
PT	comprises detecting the level of reagent that selectively associates with	
PT	a specified polynucleotide or polypeptide in a sample from the subject.	
XX		

PS Claim 1; SEQ ID NO 3; 85bp; English.

CC This invention relates to a novel method to determine whether a subject
CC has or is predisposed to a mental or neuropsychiatric disorder.
CC Specifically, it refers to identifying modulators of two genes expressed
CC in the central nervous system of patients suffering from bipolar disorder
CC or schizophrenia, namely alpha-type II calcium/calmodulin dependent
CC protein kinase (CAMKII-alpha) and TBR1 (a putative transcription factor
CC related to the Brachyury gene). The present invention describes screening
CC assays used to identify gene expression modulators, as well as
CC appropriate antibodies, agonists and antagonists thereof. Accordingly,
CC these compositions, which exhibit neuroleptic and antidepressant
CC activities can also be used for gene therapy purposes to treat the
CC aforementioned conditions, and mental disorders including mood disorders,
CC psychosis and major depression. This polynucleotide sequence is the human
CC TBR1 DNA of the invention.

XX Sequence 3186 BP; 692 A; 998 C; 774 G; 722 T; 0 U; 0 Other;

Query Match 100.0%; Score 3186; DB 12; Length 3186;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 3186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGTGATTATCTTAATATGCTATCTAATTAATTACTGTGACAGCTAACCAATGG 60
DB 1 CAGGTGATTATCTTAATATGCTATCTAATTAATTACTGTGACAGCTAACCAATGG 60
QY 61 CAGGAGCCGTTTCAATCGGCTGCAACAAGCAAGATCAAAAGTGACCTTTTCTGATTGC 120
DB 61 CAGGAGCCGTTTCAATCGGCTGCAACAAGCAAGATCAAAAGTGACCTTTTCTGATTGC 120
QY 121 TGCATAGTGCATTTGGGCAATCTCTTCCAGGGGAAAAAAGTAAATCAAACTT 180
DB 121 TGCATAGTGCATTTGGGCAATCTCTTCCAGGGGAAAAAAGTAAATCAAACTT 180
QY 181 TGAGAGCAATTTGCTGGTGAAGTCTTCTGTCTAGTGAGGGGGTCTGTGATTTCTAG 240
DB 181 TGAGAGCAATTTGCTGGTGAAGTCTTCTGTCTAGTGAGGGGGTCTGTGATTTCTAG 240
QY 241 TTTTATGATTAATAGACTTTAAAAACAGAGACGGGAGCGAGTTCAGATTCTAAG 300
DB 241 TTTTATGATTAATAGACTTTAAAAACAGAGACGGGAGCGAGTTCAGATTCTAAG 300
QY 301 CTATGAGCTGGAGCACTGCTTTCTCTTATCATGCTCTCCAAAGAAATTTCTCAATG 360
DB 301 CTATGAGCTGGAGCACTGCTTTCTCTTATCATGCTCTCCAAAGAAATTTCTCAATG 360
QY 361 TGAGCAGCAGCTAACCCACATTCAGCGGATCCGAGCTTGTCTTGACATATCCCATTA 420
DB 361 TGAGCAGCAGCTAACCCACATTCAGCGGATCCGAGCTTGTCTTGACATATCCCATTA 420
QY 421 TCTTCGACCACTGACACTTGAGAGAGATTCACTTTGAAAAAATTTACAGGGGGATGA 480
DB 421 TCTTCGACCACTGACACTTGAGAGAGATTCACTTTGAAAAAATTTACAGGGGGATGA 480
QY 481 CGAATAGTCAGATACAGCAATTTTCTGAGCTCCAGAGCTCAACAGGGGAGCTCCAGA 540
DB 481 CGAATAGTCAGATACAGCAATTTTCTGAGCTCCAGAGCTCAACAGGGGAGCTCCAGA 540
QY 541 GAAGTAACTCTCTCTCTTGAAGCGGGGTCTTGAGCTTGTGACAGATTTCATGAGCT 600
DB 541 GAAGTAACTCTCTCTCTTGAAGCGGGGTCTTGAGCTTGTGACAGATTTCATGAGCT 600
QY 601 CTGCTGCAATGCTACTCTCTCTGAGAGGGGTCTGAGCTTGTGACAGATTTCATGAGCT 660
DB 601 CTGCTGCAATGCTACTCTCTCTGAGAGGGGTCTGAGCTTGTGACAGATTTCATGAGCT 660
QY 661 CGAGTGCAGATGCTACTCTCTCTGAGAGGGGTCTGAGCTTGTGACAGATTTCATGAGCT 720
DB 661 CGAGTGCAGATGCTACTCTCTCTGAGAGGGGTCTGAGCTTGTGACAGATTTCATGAGCT 720
QY 721 GGAGCCTTAGCGCTACATGCGCCACACCGGCTCATCAACGAGAGCTTCAACAGCC 780
DB 721 GGAGCCTTAGCGCTACATGCGCCACACCGGCTCATCAACGAGAGCTTCAACAGCC 780

DB 721 GCAGCCCTTAGCGCTACATGCGCCACACCGGCTCATCAACGAGAGCTTCAACAGCC 780
QY 781 TCCGTGCAACTCTCTGCGCGAGGGATATCCCAACGGCGGGCTACCCCTACCCACAGAGT 840
DB 781 TCCGTGCAACTCTCTGCGCGAGGGATATCCCAACGGCGGGCTACCCCTACCCACAGAGT 840
QY 841 ACGGCCTCTCTACCAAGAGAGCTCCGTTTACAGATTCTCTCCACCAGCCGGGGCTGG 900
DB 841 ACGGCCTCTCTACCAAGAGAGCTCCGTTTACAGATTCTCTCCACCAGCCGGGGCTGG 900
QY 901 TGCCCGGAAAGCAGAGTGTACTGTGCAACAGGCCCTTTGGCTCAAAATTTCAACGGC 960
DB 901 TGCCCGGAAAGCAGAGTGTACTGTGCAACAGGCCCTTTGGCTCAAAATTTCAACGGC 960
QY 961 ACCAAAGGAGATGATCATCAACCAAGGGAAGCGCATGTTTCTTTTAAATTTTA 1020
DB 961 ACCAAAGGAGATGATCATCAACCAAGGGAAGCGCATGTTTCTTTTAAATTTTA 1020
QY 1021 ACATTTCTGCTGATCCCAAGGCTGATTAACAATATTTTGTGATGTGATTTTGGCGG 1080
DB 1021 ACATTTCTGCTGATCCCAAGGCTGATTAACAATATTTTGTGATGTGATTTTGGCGG 1080
QY 1081 ATCCCAATCACTGAGAGTTTCAAGAGGCAATGGGTTCTTGCGGCAAAAGCGACCA 1140
DB 1081 ATCCCAATCACTGAGAGTTTCAAGAGGCAATGGGTTCTTGCGGCAAAAGCGACCA 1140
QY 1141 ATGTGCAAGGAAATCGGGGTCTATGATCCGATCCGATCCCACTGCGGCTCACTGGA 1200
DB 1141 ATGTGCAAGGAAATCGGGGTCTATGATCCGATCCGATCCCACTGCGGCTCACTGGA 1200
QY 1201 TGCGCCAAAGAAATCTTTTGAATAATTAACCTTACGAAACAAAGAGCTTCAATA 1260
DB 1201 TGCGCCAAAGAAATCTTTTGAATAATTAACCTTACGAAACAAAGAGCTTCAATA 1260
QY 1261 ACAATGGCAGATGGTGTATTACATGCTTTCGACAAAGTACAGCCCGCTGATGTGG 1320
DB 1261 ACAATGGCAGATGGTGTATTACATGCTTTCGACAAAGTACAGCCCGCTGATGTGG 1320
QY 1321 TGGAAGTGAAGAGAGACGGGACGGAGGACACTAGCCAGCCCGGCGGTGAGAGTTCA 1380
DB 1321 TGGAAGTGAAGAGAGACGGGACGGAGGACACTAGCCAGCCCGGCGGTGAGAGTTCA 1380
QY 1381 CTTTCCCTGAGACTCAGTTCAATCGCGCTACCGGCTTACAGAAACGAGATTTACACAC 1440
DB 1381 CTTTCCCTGAGACTCAGTTCAATCGCGCTACCGGCTTACAGAAACGAGATTTACACAC 1440
QY 1441 TGAATAATGATCAACACCTTTTGCAAAAGATTTGGGATTAATTATGACAGATCTTACA 1500
DB 1441 TGAATAATGATCAACACCTTTTGCAAAAGATTTGGGATTAATTATGACAGATCTTACA 1500
QY 1501 CCGGCTGTGACATGAGACCGGCTGACCCGCTGGCCCAAGATGCGCGGCTCGAGATCG 1560
DB 1501 CCGGCTGTGACATGAGACCGGCTGACCCGCTGGCCCAAGATGCGCGGCTCGAGATCG 1560
QY 1561 TGCCGGGGCCCGGTACGCGCATGCGCGGCTTTTCTGAGAGGACAGATGCTGAGCAACT 1620
DB 1561 TGCCGGGGCCCGGTACGCGCATGCGCGGCTTTTCTGAGAGGACAGATGCTGAGCAACT 1620
QY 1621 ACGCCAGGCGCGCTTCCACCCGGGCGGGCGGGCCCGGGCTCGGATACGACCGCA 1680
DB 1621 ACGCCAGGCGCGCTTCCACCCGGGCGGGCGGGCCCGGGCTCGGATACGACCGCA 1680
QY 1681 GCGTGCAGCAACCAAGGGGTGCTGTGCGCGGACAGAGCGGAGACCGGGGCGCGCT 1740
DB 1681 GCGTGCAGCAACCAAGGGGTGCTGTGCGCGGACAGAGCGGAGACCGGGGCGCGCT 1740
QY 1741 CGCGGCAACGCTGTTTGTGACGCGGCAACCGGCTTGAATTTGCGGCTTCTGAGCT 1800
DB 1741 CGCGGCAACGCTGTTTGTGACGCGGCAACCGGCTTGAATTTGCGGCTTCTGAGCT 1800
QY 1801 ATGACACGGCCACGAGCTTCCGGGCAACCGGCTTGAATTTGCGGCTTCTGAGCT 1860
DB 1801 ATGACACGGCCACGAGCTTCCGGGCAACCGGCTTGAATTTGCGGCTTCTGAGCT 1860

QY 1861 GGTGAAGCGCTGCGCTGAGGCTGCAAGCTGCACTGCGCGCCGCTGGCTACTACG 1920
 DB 1861 GGTGAAGCGCTGCGCTGAGGCTGCAAGCTGCAAGCTGCGCGCCGCTGGCTACTACG 1920
 QY 1921 CGGACCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
 DB 1921 CGGACCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
 QY 1981 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
 DB 1981 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
 QY 2041 CCTACCTGGGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2100
 DB 2041 CCTACCTGGGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2100
 QY 2101 CGGCGGAGAGCT 2160
 DB 2101 CGGCGGAGAGCT 2160
 QY 2161 CGATCAAGTCCATGCACTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCA 2220
 DB 2161 CGATCAAGTCCATGCACTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCA 2220
 QY 2221 TCTCG 2280
 DB 2221 TCTCG 2280
 QY 2281 CCCAGCGGAGCTGCGGAGAGCTGCGGAGAGCTGCGGAGAGCTGCGGAGAGCTGCGG 2340
 DB 2281 CCCAGCGGAGCTGCGGAGAGCTGCGGAGAGCTGCGGAGAGCTGCGGAGAGCTGCGG 2340
 QY 2341 CGCAGAGTGAAGCG 2400
 DB 2341 CGCAGAGTGAAGCG 2400
 QY 2401 CCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 DB 2401 CCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 QY 2461 ATTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 DB 2461 ATTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 QY 2521 CTTTGTGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
 DB 2521 CTTTGTGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
 QY 2581 TTGGTTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 2640
 DB 2581 TTGGTTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 2640
 QY 2641 GTTCTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 2700
 DB 2641 GTTCTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 2700
 QY 2701 TGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
 DB 2701 TGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
 QY 2761 CATAGAGTGTGACTGTGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 2820
 DB 2761 CATAGAGTGTGACTGTGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 2820
 QY 2821 GATGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGT 2880
 DB 2821 GATGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGT 2880
 QY 2881 ATGATGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGT 2940
 DB 2881 ATGATGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGT 2940

QY 2941 ACAAAACAGTTCATGCTTAACCTTTTTCCTTTCTTTCTTTCTTTCTTTCTTTCTTC 3000
 DB- 2941 ACAAAACAGTTCATGCTTAACCTTTTTCCTTTCTTTCTTTCTTTCTTTCTTTCTTC 3000
 QY 3001 TCTCATACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3060
 DB 3001 TCTCATACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3060
 QY 3061 TAGAAACATGAATATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3120
 DB 3061 TAGAAACATGAATATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3120
 QY 3121 AATATTACTTGTGCGCTTAATGACACAAATAGCTAAGAGATTCACCCAAACCTT 3180
 DB 3121 AATATTACTTGTGCGCTTAATGACACAAATAGCTAAGAGATTCACCCAAACCTT 3180
 QY 3181 TAAAG 3186
 DB 3181 TAAAG 3186

RESULT 2
 AEA79130
 ID AEA79130 standard; cDNA: 1793 BP.
 XX
 AC AEA79130;
 XX
 DT 11-AUG-2005 (first entry)
 XX
 DE Human apoptosis-associated cDNA SEQ ID 378.
 XX
 KW apoptosis; cell death; cytosolic; neuroprotective; immunosuppressive;
 KW antineuritic; antitumor; dermatological; antineuritic; antineuritic;
 KW heparin; virus; neurotropic; antineuritic; antineuritic; antineuritic;
 KW vasotrophic; cerebroprotective; antineuritic; antineuritic; antineuritic;
 KW autoimmune disease; degenerative disease; viral infection; leukemia;
 KW carcinoma; sarcoma; multiple sclerosis; rheumatoid arthritis; diabetes;
 KW lupus; hepatitis; influenza viruses; Alzheimer's disease;
 KW Huntington's disease; Parkinson's disease; reperfusion injury; stroke;
 KW alcoholic liver disease; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003058021-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 13-JAN-2003; 2003WO-BP000270.
 XX
 PR 11-JAN-2002; 2002DE-01000856.
 XX
 PA (XANT-) XANTOS BIOMEDICINE AG.
 XX
 PI Koenig-Hoffman K, Kazinski M, Schaefer R, Kesper B;
 XX
 DR WPI; 2003-542134/51.
 XX
 PT New nucleic acids involved in apoptosis, useful for diagnosis and
 PT treatment of e.g. tumors and degenerative disease, also related proteins,
 PT antibodies and modulators.
 XX
 PS Claim 1a; SEQ ID NO 378; 517pp; German.
 XX
 CC This invention describes novel nucleic acid molecules that are associated
 CC with apoptosis and encode a polypeptide and are derived from a normalised
 CC gene library (embryonic or liver) or clone collections, and the extent of
 CC apoptosis measured by cell death detection assay or the CPRG assay
 CC (measuring loss of membrane integrity). The products of the invention
 CC have cytosolic, neuroprotective, immunosuppressive, antineuritic,
 CC antitumor, dermatological, antineuritic, antineuritic, antineuritic,
 CC neurotropic, antineuritic, antineuritic, antineuritic, antineuritic,
 CC cerebroprotective and antineuritic activity and can be used for gene
 CC therapy. The polynucleotides also related vectors, hosts (or their

OY	225	GCCTGGAATTCCTGATTTATGATAAATGGACTTTAAATAACAGGAGCGGAGCGGACG	284
Db	1	GTCTGGAGTTTCTGATTTATGATTAATATGGACTTTAAATAACAGGAGCGGAGCGGACG	60
OY	285	TGTTCAAGGTTCTAGAGCTATGACGCTGAGCACTGCTTCTCTTCTATCATGCTCTCC	344
Db	61	TGTTCAAGGTTCTAGAGCTATGACGCTGAGCACTGCTTCTCTTCTATCATGCTCTCC	120
OY	345	AAGAAATTTCTCAATGTGAGCGACGACTACCAATTCAGGCGGATCCGACTTGTCTTG	404
Db	121	AAGAAATTTCTCAATGTGAGCGACGACTACCAATTCAGGCGGATCCGACTTGTCTTG	180
OY	405	CACGATCATCCCATTTATCTCGACCACTGCAACTCGAGAGAAAGTTCACTTTGAAAAAA	464
Db	181	CACGATCATCCCATTTATCTCGACCACTGCAACTCGAGAGAAAGTTCACTTTGAAAAAA	240
OY	465	ATTACCAAGGGGATGACGATCAGTCAGATACAGACAATTTTCTGACTCCAAAGACTCA	524
Db	241	ATTACCAAGGGGATGACGATCAGTCAGATACAGACAATTTTCTGACTCCAAAGACTCA	300
OY	525	CCAGGGGACGTCACAGAAATTAACCTCTCTCTGCTTGGAGAGGGGTCCTAGACTTGT	584
Db	301	CCAGGGGAGTCACAGAAATTAACCTCTCTCTGCTTGGAGAGGGGTCCTAGACTTGT	360
OY	585	CACAGTTTCGATGGGCTCTGCTCAGATCGCTACCTCTCTCAGTCGACGACGACAG	644
Db	361	CACAGTTTCGATGGGCTCTGCTCAGATCGCTACCTCTCTCTCAGTCGACGACGACAG	420
OY	645	TCTGGGCGCATCTGCTCCAGTGCATGTTCCCGTACCCCGGCGAGCAAGAACCGGCGAC	704
Db	421	TCTGGGCGCATCTGCTCCAGTGCATGTTCCCGTACCCCGGCGAGCAAGAACCGGCGAC	480
OY	705	CCCGGCTTCTCATGCGGCGACCCCTAGCGGCTACATGGCCCAACAACCGGCTCATCAAC	764
Db	481	CCCGGCTTCTCATGCGGCGACCCCTAGCGGCTACATGGCCCAACAACCGGCTCATCAAC	540
OY	765	GGAGGCTTCAACAGCCTCTGTGCCAATCTCTGCGCGACAGGATACCCCAACGGCGGCTAC	824
Db	541	GGAGGCTTCAACAGCCTCTGTGCCAATCTCTGCGCGACAGGATACCCCAACGGCGGCTAC	600
OY	825	CCCTTACCCACAGCAGTACGGGCACTCTTACCAAAGAGCTCCGTTCTAACAGATTCTCTCC	884
Db	601	CCCTTACCCACAGCAGTACGGGCACTCTTACCAAAGAGCTCCGTTCTAACAGATTCTCTCC	660
OY	885	ACCCAGCGGGGCTGGTGCCTCGGCAAAAGACAGGTATCCTGTGCAACAGGCCCTTTGG	944
Db	661	ACCCAGCGGGGCTGGTGCCTCGGCAAAAGACAGGTATCCTGTGCAACAGGCCCTTTGG	720
OY	945	CTGAATATTCACCGGACCAAAACGAGATGATCATACCAAAACAGGGAAGCGCATGTT	1000
Db	721	CTGAATATTCACCGGACCAAAACGAGATGATCATACCAAAACAGGGAAGCGCATGTT	780
OY	1005	CCTTTTTTAAGTTTAACTTTCTGGTCTCGATCCACGGGCTCATTAACAATATTTTGG	1060
Db	781	CCTTTTTTAAGTTTAACTTTCTGGTCTCGATCCACGGGCTCATTAACAATATTTTGG	840

DT 23-JAN-2002 (first entry)

XX Human nervous system related polynucleotide SEQ ID NO 6913.
DE Human; neurotropic; neuroprotective; cytostatic; dermatological; virocidic;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisking; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antituber; anticonvulsant; antitungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; Gene therapy; vaccine; ds.
OS Homo sapiens.
XX WO200159063-A2.
XX 16-AUG-2001.
PD 17-JAN-2001; 2001WO-US001334.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232357P.
PR 14-SEP-2000; 2000US-0232358P.

PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235833P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241825P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251909P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Disclosure; SEQ ID NO 6913; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pcr_sequences
XX
XX
SQ Sequence 6878 BP; 2032 A; 1487 C; 1634 G; 1725 T; 0 U; 0 Other;
Query Match 53.2%; Score 1694; DB 5; Length 6878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1493 GATCTACACCGGCTGTGATGACCGCTGACCCCTTCGCCCAACGACTGCGCGCTC 1552
DB 2504 GATCTACACCGGCTGTGATGACCGCTGACCCCTTCGCCCAACGACTGCGCGCTC 2445
QY 1553 GCAGATCGTGGCCCGGGGCGCCGCTACCGCATGCGCGCTCTTCTCTGCAAGAACATTCGT 1612
DB 2444 GCAGATCGTGGCCCGGGGCGCCGCTACCGCATGCGCGCTCTTCTCTGCAAGAACATTCGT 2385
QY 1613 GAGCACTACGCAAGGCGCCGCTTCCACCGGGGCGGGGCGGGCCCGGGGCTAC 1672
DB 2384 GAGCACTACGCAAGGCGCCGCTTCCACCGGGGCGGGGCGGGCCCGGGGCTAC 2325
QY 1673 GAGCGCAGAGTGCCTGCAACCAAGGCGCTGTCTGCGCGAGCAGAGCGCGAGACCCGGG 1732
DB 2324 GAGCGCAGAGTGCCTGCAACCAAGGCGCTGTCTGCGCGAGCAGAGCGCGAGACCCGGG 2265
QY 1733 CGGCGCTTCGCGCAACGCTGTGTTGTGACGCGCGCAACCGGCTGACTTCGCGGC 1792
DB 2264 CGGCGCTTCGCGCAACGCTGTGTTGTGACGCGCGCAACCGGCTGACTTCGCGGC 2205
QY 1793 CTGCGGCTATGACACGCGCAACGACTTCGCGGGCAACGCGCAACGCTCTCTTAACG 1852
DB 2204 CTGCGGCTATGACACGCGCAACGACTTCGCGGGCAACGCGCAACGCTCTCTTAACG 2145
QY 1853 GCGGCGGCGGCGTGAAGCGCTGCGCTGACGCTGACGAGCTGACCTGCGCGCGCTCGG 1912
DB 2144 GCGGCGGCGGCGTGAAGCGCTGCGCTGACGAGCTGACGAGCTGCGCGCGCGCTCGG 2085

QY 1913 CTACTACGCGACCCGTCGCGGCTGGGGGCGCCCGCAGTCCCGCGAGTACTGCGGCAACAA 1972
DB 2084 CTACTACGCGACCCGTCGCGGCTGGGGGCGCCCGCAGTCCCGCGAGTACTGCGGCAACAA 2025
QY 1973 GTCGCGCTCGTGTGCTGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGCGCATGCGCG 2032
DB 2024 GTCGCGCTCGTGTGCTGCTGCTGCGCCCAACAGCGCGCGCGCGCGCGCGCATGCGCG 1965
QY 2033 CGCCATTCCTTACTCTGCGCGAGAGGCGCGAGGCGCTGCGCGCGCGCGCGCGCGCTGCC 2092
DB 1964 CGCCATTCCTTACTCTGCGCGAGAGGCGCGAGGCGCTGCGCGCGCGCGCGCGCGCTGCC 1905
QY 2093 GCCCGGCGCGCGCGAGAGCGCGCAAGCCCAAGACCTGTCCGACTTCAGTGCATGAGAC 2152
DB 1904 GCCCGGCGCGCGCGAGAGCGCGCAAGCCCAAGACCTGTCCGACTTCAGTGCATGAGAC 1845
QY 2153 GCCCTCTCGATCAAGTCCATCGACTCGCGGATTTTACGACGAGCGCAAGCG 2212
DB 1844 GCCCTCTCGATCAAGTCCATCGACTCGCGGATTTTACGACGAGCGCAAGCG 1785
QY 2213 GAGCGGATCTGCGCGCGCGCGCACGCGCGCTGCCGAGATTCGCCCGCTCAAGACGA 2272
DB 1784 GAGCGGATCTGCGCGCGCGCGCACGCGCGCTGCCGAGATTCGCCCGCTCAAGACGA 1725
QY 2273 GGTGCTGCGCGAGCGGAGCTGCGAAGAACTGCGCGCAAGACATTAAGCGCTACTATG 2332
DB 1724 GGTGCTGCGCGAGCGGAGCTGCGAAGAACTGCGCGCAAGACATTAAGCGCTACTATG 1665
QY 2333 CTCTACTCGACAGACTAGGCGCGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2392
DB 1664 CTCTACTCGACAGACTAGGCGCGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1605
QY 2393 AGCCAGCCCTCACAGCTCTTCCCGAGTCCGCTCCCGCACATCTCTCTTGGCCACCA 2452
DB 1604 AGCCAGCCCTCACAGCTCTTCCCGAGTCCGCTCCCGCACATCTCTCTTGGCCACCA 1545
QY 2453 CTCAATTATTTGACCCCTCGATGAGCGCGCTGACAGAAATAGTACAGGCTCCGAGCGT 2512
DB 1544 CTCAATTATTTGACCCCTCGATGAGCGCGCTGACAGAAATAGTACAGGCTCCGAGCGT 1485
QY 2513 ATTTTAACCTTTTGTGACAGACAGTCTGCAATTAAGTCAACGACTTCACTTGTG 2572
DB 1484 ATTTTAACCTTTTGTGACAGACAGTCTGCAATTAAGTCAACGACTTCACTTGTG 1425
QY 2573 TAAACCTTTGCTTCTTCTACTACTCTTCTTCTGAGAGTAACTCTTCAATTCGCC 2632
DB 1424 TAAACCTTTGCTTCTTCTACTACTCTTCTTCTGAGAGTAACTCTTCAATTCGCC 1365
QY 2633 TCCCGCTGCTCTTCTTCACTGCTTCTTCTTCTGTAAGAACTCTTCACTT 2692
DB 1364 TCCCGCTGCTCTTCTTCACTGCTTCTTCTTCTGTAAGAACTCTTCACTT 1305
QY 2693 AGAGACCTGGGAGTCTGTCAGGACAGACGATTCGACCGCGCAAGTCTGCGCTCC 2752
DB 1304 AGAGACCTGGGAGTCTGTCAGGACAGACGATTCGACCGCGCAAGTCTGCGCTCC 1245
QY 2753 ACATTAAACATGAGATGTTGACTTGAACCTGGAACCAACCAAGCGGCTCTTCTATC 2812
DB 1244 ACATTAAACATGAGATGTTGACTTGAACCTGGAACCAACCAAGCGGCTCTTCTATC 1185
QY 2813 CCCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2872
DB 1184 CCCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
QY 2873 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2932
DB 1124 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
QY 2933 GTCAAGGACAAACCAAGTCAATGCTTAACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2992
DB 1064 GTCAAGGACAAACCAAGTCAATGCTTAACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1005


```
DB 185 CTACCACTTCGAGCGGATCCGAGCTGTCTGTGACATCATCCCATTTATCTGACCA 244
OY 431 TGACAACTCGAGAGAGAGTTCACTTTGAAAAAATTAACGAGGGAGTACGATCAGTC 490
DB 245 TGACAACTCGAGAGAGAGTTCACTTTGAAAAAATTAACGAGGGAGTACGATCAGTC 304
OY 431 AGATCAGACAAATTTTCTGACTCGAAGACTCACCGAGGAGCTCCAGAGAGTAACACT 550
DB 305 AGATCAGACAAATTTTCTGACTCGAAGACTCACCGAGGAGCTCCAGAGAGTAACACT 364
OY 551 CTCTCTGTCTTGAAGGAGGCTCTC 574
DB 365 CTCTCTGTCTTGAAGGAGGCTCTC 388

RESULT 8
ACH15279
ID ACH15279 standard; cDNA; 457 BP.
XX
AC ACH15279;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult brain cDNA #2491.
XX
KM Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 2491; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics or for physical mapping of the human genome,
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
```

```
XX
SQ Sequence 457 BP; 81 A; 168 C; 85 G; 123 T; 0 U; 0 Other:
XX
Query Match 9.5%; Score 303; DB 9; Length 457;
Best Local Similarity 99.7%; Pred. No. 9.1e-116;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2372 CCGCGCGGGGCGGAGCCCCCAGCAGCCCTCAGACTCTTCCCGAGTCCGCTCC 2431
DB 104 CCGCGCGGGGCGGAGCCCCCAGCAGCCCTCAGACTCTTCCCGAGTCCGCTCC 163
OY 2432 ACATCTCTCTTGGCAGCCCACTCATTTTAATTTGACCTCGATGGCGGTGACGGAAT 2491
DB 164 ACATCTCTCTTGGCAGCCCACTCATTTTAATTTGACCTCGATGGCGGTGACGGAAT 223
OY 2492 AAGTGCAGGTCTCCGAGCGATTTTAACCTTTTGGCAGAGAGTCTTGCAATTAGCT 2551
DB 224 AAGTGCAGGTCTCCGAGCGATTTTAACCTTTTGGCAGAGAGTCTTGCAATTAGCT 283
OY 2552 CACGACCTTCAACTTTGTGTAAACCTTTGTTTGTTTCTTACTTCTTCTGTGGA 2611
DB 284 CACGACCTTCAACTTTGTGTAAACCTTTGTTTGTTTCTTACTTCTTCTGTGGA 343
OY 2612 GTTATCTCTTCAATTCCTCCCTCCCTCGTCTTCTTCTTAACCTTCTTCTT 2671
DB 344 GTTATCTCTTCAATTCCTCCCTCCCTCGTCTTCTTCTTAACCTTCTTCTT 403
OY 2672 GTAATGAACCTTTCACCTTTAGAGAGACTGGGAGTCTGTGAGAGCAGCG 2725
DB 404 GGAATGAACCTTTCACCTTTAGAGAGACTGGGAGTCTGTGAGAGCAGCG 457

RESULT 9
ACF88226
ID ACF88226 standard; DNA; 463 BP.
XX
AC ACF88226;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human SIRS/sepsis diagnostic marker DNA fragment 7086.
XX
KM Systemic inflammatory response syndrome; SIRS; antibacterial;
XX immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
OS Homo sapiens.
XX
PN WO2004087949-A2.
XX
PD 14-OCT-2004.
XX
PF 31-MAR-2004; 2004WO-EP003419.
XX
PR 02-APR-2003; 2003DE-01015031.
PR 08-AUG-2003; 2003DE-01036511.
PR 02-SEP-2003; 2003DE-01040395.
XX
PA (SIRS-) SIRS LAB GMBH.
XX
PI Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Daigner H;
XX
DR WPI; 2004-748070/73.
XX
PT In vitro detection of systemic inflammatory response syndrome and related
PT conditions, for e.g. monitoring progression, comprises detecting abnormal
PT expression of disease-related genes.
XX
PS Disclosure; Page: 75pp; German.
XX
CC The invention relates to a novel method for in vitro detection of
CC systemic inflammatory response syndrome (SIRS). The method comprises
CC detecting abnormal expression of disease-related genes, or their
CC associated peptides. The method of the invention demonstrates
```

antibacterial, immunosuppressive and antiinflammatory applications and
 CC may be used for early differential diagnosis, monitoring progression,
 CC assessing risk, assessing the likely response to treatment and for post
 CC mortem diagnosis of systemic inflammatory response syndrome, sepsis and
 CC sepsis-like conditions. The recombinant or synthetic nucleic acid
 CC sequences of the invention, or derived proteins or peptides may be
 CC useful as calibrants in assays for the specified diseases, for evaluating
 CC activity or toxicity in screening for active agents and/or for
 CC preparation of agents for treatment or prevention of the specified
 CC diseases. The current sequence is that of a human SIRS/sepsis diagnostic
 CC marker DNA fragment of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at ftp.wipo.int/pub/published
 CC pct sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
 CC disclosed within the specification, however, these have not been taken
 CC into account during indexing due to inconsistencies in application and
 CC format

Sequence 463 BP; 86 A; 139 C; 97 G; 137 T; 0 U; 4 Other;

Query Match 8.8%; Score 281; DB 13; Length 463;
 Best Local Similarity 100.0%; Pred. No. 4.2e-125;
 Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2415 CCCAGCTCCGCTCCGACACTCTCTTGGGACCCACTCATTTATTGACCTCGAT 2474
 1 CCCAGCTCCGCTCCGACACTCTCTTGGGACCCACTCATTTATTGACCTCGAT 60
 2475 GGGCGCTGCGAGCAATAGTCAAGTCTCCGAGCGTGATTTTAACTTTTGACAGC 2534
 61 GGGCGCTGCGAGCAATAGTCAAGTCTCCGAGCGTGATTTTAACTTTTGACAGC 120
 2535 AGCTCTGCAATTAGTCAAGCACTTCACTTGTGTAACTTTTGCTTCTACT 2594
 121 AGCTCTGCAATTAGTCAAGCACTTCACTTGTGTAACTTTTGCTTCTACT 180
 2595 TACTCTTCTGTGTGAGTTATCTCTCAATTTCCCTGCTCTTCTCTTAC 2654
 181 TACTCTTCTGTGTGAGTTATCTCTCAATTTCCCTGCTCTTCTCTTAC 240
 2655 TCTCTACTTCTCTTCTTGTAACTTCACTTCACTTAC 2695
 241 TCTCTACTTCTCTTCTTGTAACTTCACTTCACTTAC 281

RESULT 10

AAQ39695
 ID AAQ39695 standard; DNA; 389 BP.

AAQ39695;

25-MAR-2003 (revised)
 20-MAY-1993 (first entry)

Expressed Sequence Tag human gene marker EST00360.

expressed sequence tag; human genome project; chromosome;
 human gene sequencing; PCR mapping; somatic cell hybrids;
 sublocalisation; gene tagging; tissue typing.

Synthetic.

WO9300353-A1.

07-JAN-1993.

19-JUN-1992; 92WO-US005222.

20-JUN-1991; 91US-00716831.

12-FEB-1992; 92US-00837195.

(USSH) US DEPT HEALTH & HUMAN SERVICE.

Venter JC, Adams MD;
 WPI, 1993-036325/04.
 Particular expressed sequence tags from human CDNA - corresponds to
 PT transcrip prod.; of genes, useful for tagging genes, mapping
 PT chromosomes and tissue typing.

Claim 3; Page 105; 199pp; English.

This sequence represents an EST (expressed sequence tag) ESTs are markers
 CC for human genes actually transcribed in vivo. Unlike the random genomic
 CC DNA sequence tagged sites (STS), ESTs point directly to expressed genes.
 CC The use of ESTs could facilitate the tagging of most expressed human
 CC genes within a few years at a fraction of the cost of complete genomic
 CC sequencing. Using PCR primers AAQ39419-039580 (sequences designed from
 CC the ESTs) sublocalisation of an EST can be achieved with panels of
 CC fragments from specific chromosomes or pools of large genomic clones in
 CC an analogous manner. This sequence represents EST00360. (Updated on 25-
 CC MAR-2003 to correct PN field.)

Sequence 389 BP; 97 A; 81 C; 71 G; 139 T; 0 U; 1 Other;

Query Match 8.2%; Score 260; DB 2; Length 389;
 Best Local Similarity 99.7%; Pred. No. 6.3e-115;
 Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2775 TCTAGAACCTGACCCACCCAGCGCTCTTCTTATCCCGAGTGAATGATGATGA 2834
 1 TCTAGAACCTGACCCACCCAGCGCTCTTCTTATCCCGAGTGAATGATGATGA 60
 2835 TGAATGAGGAGTGAATTAATTTAGTGAACAAGCTGTGAATGATGATGATGA 2894
 61 TGAATGAGGAGTGAATTAATTTAGTGAACAAGCTGTGAATGATGATGATGA 120
 2895 TGTATATTATTTAGTGAAGATGCTATGTTTATTTCTGTCAGAGCACAAAACAGTTCA 2954
 121 TGTATATTATTTAGTGAAGATGCTATGTTTATTTCTGTCAGAGCACAAAACAGTTCA 180
 2955 TGTATATTATTTAGTGAAGATGCTATGTTTATTTCTGTCAGAGCACAAAACAGTTCA 3014
 181 TGTATATTATTTAGTGAAGATGCTATGTTTATTTCTGTCAGAGCACAAAACAGTTCA 240
 3015 TTTCTCTCTTTTAAATTTCTTGTGAGTAAATATCTTAAGAGGCTCTGAACATGAAT 3074
 241 TTTCTCTCTTTTAAATTTCTTGTGAGTAAATATCTTAAGAGGCTCTGAACATGAAT 300

RESULT 11

AAQ59107
 ID AAQ59107 standard; cDNA; 389 BP.

AAQ59107;

25-MAR-2003 (revised)
 16-MAR-1994 (first entry)

Human brain Expressed Sequence Tag EST00360.

Gene transcription product; genetic markers; tagging; in vivo;
 transcripion; mapping; locations; chromosomes; chromosomal; ss.

Homo sapiens.

WO9316178-A2.

19-AUG-1993.

12-FEB-1993; 93WO-US001294.

```
XX 12-FEB-1992; 92US-00837195.
PR (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX Venter CJ, Adams MD, Moreno RF;
XX WPI; 1993-272882/34.
XX Enriched oligonucleotides and corresp. sequences - used as markers for
PT human genes transcribed in-vivo, facilitate tagging of most human genes.
XX Example 1; Page 122; 500pp; English.
XX The Expressed Sequence Tag was isolated from a human brain cDNA library
CC as part of a large set of ESTs which can be used as markers for human
CC genes transcribed in vivo. They can be used to facilitate tagging of most
CC human genes, for mapping locations of expressed genes on chromosomes, for
CC individual or forensic identification, for mapping locations of disease-
CC associated genes, for identification of tissue type, and for prepn. of
CC antisense sequences, probes and constructs. EST00360 has a "poor" coding
CC probability as evaluated using the coding-region prediction program CRM.
CC See also AA059041-Q61440. (Updated on 25-MAR-2003 to correct PN field.)
SQ Sequence 369 BP; 97 A; 81 C; 71 G; 139 T; 0 U; 1 Other;
Query Match 8.2%; Score 260; DB 2; Length 389;
Best Local Similarity 99.7%; Pred. No. 6.3e-115;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2775 TCTAGAACCTGGACCCACCACGCGCTCTTCTTATCCCGAGTGATGATGATGGA 2834
DB 1 TCTAGAACCTGGACCCACCACGCGCTCTTCTTATCCCGAGTGATGATGATGGA 60
QY 2835 TGGATGTGATGGATGTTAATTTTATGTGAACAAAGCTGTGAATGATGATGATG 2894
DB 61 TGGATGTGATGGATGTTAATTTTATGTGAACAAAGCTGTGAATGATGATGATG 120
QY 2895 TCTTAAATTTATGTGAACGAATGCTAGTTTATCTCGTCAAGGACAAACCAAGTTCA 2954
DB 121 TGTTAATTTATGTGAACGAATGCTAGTTTATCTCGTCAAGGACAAACCAAGTTCA 180
QY 2955 TGCCTAACCTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 3014
DB 181 TGCCTAACCTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 240
QY 3015 TTTCTCTCTTTTAAATTTTCTGTGAGATTAATTTCTAAGAGCTCTAAGAAACATGAAT 3074
DB 241 TTTCTCTCTTTTAAATTTTCTGTGAGATTAATTTCTAAGAGCTCTAAGAAACATGAAT 300
QY 3075 ACTCAGTAGTG 3085
DB 301 ACTCAGTAGTG 311
RESULT 12
AAC18168
ID AAC18168 standard; cDNA; 235 BP.
XX AAC18168;
AC 06-OCT-2000 (first entry)
XX 06-OCT-2000 (first entry)
DE Human secreted protein 5' EST, SEQ ID NO: 22243.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
OS Homo sapiens.
XX EPI033401-A2.
XX EPI033401-A2.
XX 06-SEP-2000.
PD 06-SEP-2000.
```

```
XX 21-FEB-2000; 2000EP-00200610.
PR 26-FEB-1999; 99US-0122487P.
XX (GEST ) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Claim 1; SEQ ID NO 22243; 71pp + Sequence Listing; English.
PS The present sequence is one of a large number of 5' ESTs derived from
XX cDNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
SQ Sequence 235 BP; 56 A; 57 C; 27 G; 95 T; 0 U; 0 Other;
Query Match 7.4%; Score 235; DB 3; Length 235;
Best Local Similarity 100.0%; Pred. No. 8.2e-103;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2293 TTTTATCTCGTCAAGGACAAACCAAGTTCAATGCTTAACTTTTCTTCTTCTTCT 2292
DB 1 TTTTATCTCGTCAAGGACAAACCAAGTTCAATGCTTAACTTTTCTTCTTCTTCTTCT 60
QY 2293 TTTGCTTTTCCTTCTCTCTCTCATACCTTCTCTTCTCTTCTTCTTCTTCTTCTTCT 3042
DB 61 TTTGCTTTTCCTTCTCTCTCTCATACCTTCTCTTCTCTTCTTCTTCTTCTTCTTCT 120
QY 3043 TAAATTTCTAAGAGGCTCTAAGAAACATGAATATCTCAGTAGTATGGTTCCACTTCT 3102
DB 121 TAAATTTCTAAGAGGCTCTAAGAAACATGAATATCTCAGTAGTATGGTTCCACTTCT 180
QY 3103 CCTCAATCCGTTGATGAATAATTAATGATGCTTAATGCAACAATAGCTA 3157
DB 181 CCTCAATCCGTTGATGAATAATTAATGATGCTTAATGCAACAATAGCTA 235
RESULT 13
ABN33794
ID ABN33794 standard; DNA; 60 BP.
XX ABN33794;
AC 15-JUL-2002 (first entry)
XX 15-JUL-2002 (first entry)
DE Human spliced transcript detection oligonucleotide SEQ ID NO:6542.
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX Homo sapiens.
OS Homo sapiens.
XX WO200210449-A2.
XX WO200210449-A2.
XX 07-FEB-2002.
PD 07-FEB-2002.
```


XX 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
XX
PA (EPIC-) EPIGENOMICS AG.
XX

PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX

PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX

PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX

CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's) and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX

SQ Sequence 1026 BP; 329 A; 429 C; 128 G; 140 T; 0 U; 0 Other;
XX

Query Match

1.0%; Score 32; DB 6; Length 1026;

Best Local Similarity 100.0%; Pred. No. 0.00019;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCTCGCCCAACGACTGCGCGCTCGCA 1555
|||||

DB 112 ACCCCTCGCCCAACGACTGCGCGCTCGCA 143
|||||

Search completed: December 20, 2005, 17:22:43
Job time : 1810 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 16:05:34 ; Search time 12030 Seconds
(without alignments)
12390.993 Million cell updates/sec

Title: US-10-649-400-3
Perfect score: 3166
Sequence: 1 cagtgatcattccatlaa.....caccacaacaccttaagg 3166

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 0

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST:
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*
10: gb_esc10:*
11: gb_esc11:*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	777	24.4	1169	3	BQ067178
2	772	24.2	820	3	BI915063
3	754	23.7	957	5	BX441549
4	741	23.3	1044	3	BM922065
5	725	22.8	1116	5	BX418703
6	701	22.0	824	5	BX418656
7	694	21.8	768	8	DR001136
8	691	21.7	921	5	BX420367
9	625	19.6	945	3	BI914132
10	611	19.2	1028	3	BM548197
11	588	18.5	872	3	BI914330
12	581	18.2	583	3	BP311133
13	536	16.8	639	3	BI819019
14	531	16.7	740	3	BI917781
15	523	16.4	831	1	AM003023
16	508	15.9	534	5	BM114402
17	501	15.7	930	3	BM807553
18	491	15.4	930	3	BI912693
19	454	14.2	582	3	BP228857
20	427	13.4	739	3	BI918597
21	426	13.4	779	3	BI820633
22	424	13.3	475	1	AM026293

C	23	419	13.2	606	1	AM590019	h29e09.x
	24	416	13.1	748	3	BI917426	BI917426
	25	415	13.0	772	3	BI823704	BI823704
	26	389	12.2	791	3	BI918360	BI918360
	27	385	12.1	583	3	BP229748	BP229748
	28	380	11.9	463	1	AL045693	AL045693
	29	359	11.3	541	1	AI201654	AI201654
	30	350	11.0	468	10	CG464729	CG464729
	31	348	10.9	966	3	BI754056	BI754056
	32	343	10.8	425	10	AY419847	AY419847
	33	335	10.5	386	1	AM160871	AM160871
	34	334	10.5	910	7	CM643299	CM643299
	35	334	10.5	910	7	CM805179	CM805179
	36	328	10.3	507	8	H41824	H41824
	37	322	10.1	507	8	RS4288	RS4288
	38	312	9.8	353	1	AM898907	AM898907
	39	312	9.8	363	1	AI815469	AI815469
	40	289	9.1	991	3	BI757271	BI757271
	41	283	8.9	386	8	D56477	D56477
	42	281	8.8	463	8	H10108	H10108
	43	275	8.6	1015	3	BM921945	BM921945
	44	260	8.2	389	8	M62287	M62287
	45	255	8.0	385	2	BF059564	BF059564

ALIGNMENTS

RESULT 1
LOCUS BQ067178 1169 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6767481 NIH_MGC_115 Homo sapiens CDNA IMAGE:5751547
ACCESSION BQ067178
VERSION BQ067178.1 GI:19896224
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
founded through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12784 row: e column: 20
High quality sequence start: 16
High quality sequence stop: 685.
Location/Qualifiers
1. 1169
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5751547"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector: pcmv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male testis, age 69; library is male lung, age 27; and 1 male testis, age 69; library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C.

FEATURES

source

Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

Query Match 24.4%; Score 777; DB 3; Length 1169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 777; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

826 CCTACCCACAGACGATACGGCCCTCTTACCAAGAGAGCTCCGTTCTACCAAGTTCTCTCCA 885
36 CCTACCCACAGACGATACGGCCCTCTTACCAAGAGAGCTCCGTTCTACCAAGTTCTCTCCA 95
886 CCCAGCCGGGGCTGGTGGCCCGCAAGACAGAGTGTACCTGTGCAACAGAGCCCTTTGGC 945
96 CCCAGCCGGGGCTGGTGGCCCGCAAGACAGAGTGTACCTGTGCAACAGAGCCCTTTGGC 155
946 TGAATTTTACCGGACCAACAGAGATGATCATCACCAACAGGGAAGCGCATGTTTC 1005
156 TGAATTTTACCGGACCAACAGAGATGATCATCACCAACAGGGAAGCGCATGTTTC 215

QY 1006 CTTTTTTAAGTTTAACTTTCTGTCTGTGATCCCAAGGCTCATTAACAATATTTTGTGG 1065
DB 216 CTTTTTTAAGTTTAACTTTCTGTCTGTGATCCCAAGGCTCATTAACAATATTTTGTGG 275

QY 1066 ATGTGATTTTGGCGGATCCCAATCACTGAGAGGTTTCAAGAGGCAAAATGGGTTCTTGG 1125
DB 276 ATGTGATTTTGGCGGATCCCAATCACTGAGAGGTTTCAAGAGGCAAAATGGGTTCTTGG 335

QY 1126 GCAAAACCGGACCAATGTGCAAGAAATGGGTTCTATATGATCCGATTCGGATTCGCCAACA 1185
DB 336 GCAAAACCGGACCAATGTGCAAGAAATGGGTTCTATATGATCCGATTCGGATTCGCCAACA 395

QY 1186 CTGGGGCTTCACTGATGAGCGCCAAAGAAATCTCTTTTGGAAATTTAAACTTACGAACAACA 1245
DB 396 CTGGGGCTTCACTGATGAGCGCCAAAGAAATCTCTTTTGGAAATTTAAACTTACGAACAACA 455

QY 1246 AAGAGCTTCAATATAAACAATGGGAGATGGTGTATTTTACAGTCTTGGCAAGTACCAAGC 1305
DB 456 AAGAGCTTCAATATAAACAATGGGAGATGGTGTATTTTACAGTCTTGGCAAGTACCAAGC 515

QY 1306 CCCGCTCGATGTGTGTAAGTGAACGAGACGGGACGGGACCACTAGCCAGCCCGGC 1365
DB 516 CCCGCTCGATGTGTGTAAGTGAACGAGACGGGACGGGACCACTAGCCAGCCCGGC 575

QY 1366 GGGTGCAGACCTTCCCTTCCCTGAGACTCACTTCACTGCGCTTCCAGCTTACCGAACA 1425
DB 576 GGGTGCAGACCTTCCCTTCCCTGAGACTCACTTCACTGCGCTTCCAGCTTACCGAACA 635

QY 1426 CGGATATTACAACTGAATAGATCAACAACCTTTTGCAAAAGGATTTCCGATAATT 1485
DB 636 CGGATATTACAACTGAATAGATCAACAACCTTTTGCAAAAGGATTTCCGATAATT 695

QY 1486 ATGACACGATCTACACCGGCTGTGACATGACCGCTGACCCCTTGGCCCAACGACTGGC 1545
DB 696 ATGACACGATCTACACCGGCTGTGACATGACCGCTGACCCCTTGGCCCAACGACTGGC 755

QY 1546 CGCGCTCGGAGATCGGCGCGCGCGCGCTAGACCAATGGCGCGCTTTTCCGACAG 1602
DB 756 CGCGCTCGGAGATCGGCGCGCGCGCGCTAGACCAATGGCGCGCTTTTCCGACAG 812

RESULT 2
BI915063 820 bp mRNA linear EST 16-OCT-2001
LOCUS 603177258F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241542 5',
DEFINITION mRNA sequence.
ACCESSION BI915063
VERSION BI915063.1 GI:16199126
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strauberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM11609 row: C column: 15
High quality sequence stop: 816.

FEATURES
source
Location/Qualifiers
1..820

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5241542"
/lab_host="DH10B"
/clone_id="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 24.2%; Score 772; DB 3; Length 820;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 GAGCGGAGGGCGAGGTGTTCAAGTCTATGACGCTGAGCACTGCTTTCTCT 329
DB 49 GAGCGGAGGGCGAGGTGTTCAAGTCTATGACGCTGAGCACTGCTTTCTCT 108

QY 330 TCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCGAGCTACCACTTACAGCGGA 389
DB 109 TCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCGAGCTACCACTTACAGCGGA 168

QY 390 TCCGAGCTTGTCTTGAAGATCATCCATTATCTGACCACTGACCACTGGAGAGT 449
DB 169 TCCGAGCTTGTCTTGAAGATCATCCATTATCTGACCACTGACCACTGGAGAGT 228

QY 450 TCACCTTTGAAAAATTTACAGGGGATGACGAATAGTCAATACAGCAATTTTCT 509
DB 229 TCACCTTTGAAAAATTTACAGGGGATGACGAATAGTCAATACAGCAATTTTCT 288

QY 510 GACTCCAGGACTACAGAGGGGAGCTCCAGAGAAATGTAACCTCTCTGTTGGAGGG 569
DB 229 GACTCCAGGACTACAGAGGGGAGCTCCAGAGAAATGTAACCTCTCTGTTGGAGGG 348

QY 289 GACTCCAGGACTACAGAGGGGAGCTCCAGAGAAATGTAACCTCTCTGTTGGAGGG 348
DB 409 TCCAGCGACCAAGTGTGCGGACCACTGCTCCAGTGCATGTTCCCGTACCCCGGACG 468

QY 570 GTCTCTGAGCTTGTGACAGTTTGTATGATGCTGTGCTGCAATGCTACTCTCTCAG 629
DB 349 GTCTCTGAGCTTGTGACAGTTTGTATGATGCTGTGCTGCAATGCTACTCTCTCAG 408

QY 630 TCCAGCGACCAAGTGTGCGGACCACTGCTCCAGTGCATGTTCCGTTACCCGGGACG 689
DB 409 TCCAGCGACCAAGTGTGCGGACCACTGCTCCAGTGCATGTTCCGTTACCCGGGACG 468

QY 690 CAGGACCGGCGACCCCGCTTCTCATGCGGACCGCTGACGCTACATGGGCCACAC 749
DB 469 CAGGACCGGCGACCCCGCTTCTCATGCGGACCGCTGACGCTACATGGGCCACAC 528

QY 750 CCGGTATCAACAAGAGCTTACAGAGCTCTGTCAACTCTCTGCGCGAGGATAC 809

```

Db      529  |||||
CGGTCATCACCAAGAGCTTACCAAGCTTCTGTCCAACTCTCGCCGAGGATAC 588
OY      810  |||||
CCACGCGCGGTACCCCTACCAAGAGCACTCTTACCAAGAGCTTCTGTC 869
Db      589  |||||
CCACGCGCGGTACCCCTACCAAGAGCACTCTTACCAAGAGCTTCTGTC 648
OY      870  |||||
TACCAATTCCTCTCAACCAAGCGGCGTGTGCTCCGCAAGCAAGTGTACTGTC 929
Db      649  |||||
TACCAATTCCTCTCAACCAAGCGGCGTGTGCTCCGCAAGCAAGTGTACTGTC 708
OY      930  |||||
AACAGAGCCCTTGTGCTGAATTTACCGGCAACCAAGAGATATCATCAACAG 989
Db      709  |||||
AACAGAGCCCTTGTGCTGAATTTACCGGCAACCAAGAGATATCATCAACAG 768
OY      990  |||||
GGAGGCGCATGTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1041
Db      769  |||||
GGAGGCGCATGTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 820

RESULT 3
BX441549 957 bp mRNA linear EST 05-MAY-2004
LOCUS     BX441549 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION
CSODF018YJ12 5-PRIME, mRNA sequence.
ACCESSION BX441549
VERSION    BX441549.2 GI:47038398
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 957)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31018481.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSODF018DE06P1c=1851.f.
location/Qualifiers
1. 957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF018YJ12"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_1lb="Homo sapiens FETAL BRAIN"
/notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 23.7%; Score 754; DB 5; Length 957;
Best local Similarity 99.9%; Pred. No. 0;
Matches 874; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY      387  |||||
GGATCCGAGCTTGTGCGATCATCATCATCATCATCATCATCATCATCATCATCAT 446
Db      8  |||||
GGATCCGAGCTTGTGCGATCATCATCATCATCATCATCATCATCATCATCATCAT 67

```

```

OY      447  |||||
AGTTACCTTTGAAAAAATTATACAGGGGATGAGATGATGATGATGATGATGATGAT 506
Db      68  |||||
AGTTACCTTTGAAAAAATTATACAGGGGATGAGATGATGATGATGATGATGATGATGAT 127
OY      507  |||||
CTGACTCCAGAGATCATCAAGAGGAGCTCCAGAGAAATTAATCTCTCTCTCTGAGAC 566
Db      128  |||||
CTGACTCCAGAGATCATCAAGAGGAGCTCCAGAGAAATTAATCTCTCTCTCTGAGAC 187
OY      567  |||||
GGGGTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 626
Db      188  |||||
GGGGTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
OY      627  |||||
CAGTCCAGGCGCAAGTCCGCGGCACTGCTCCAGATGCTGCTGCTGCTGCTGCTGCTGCT 686
Db      248  |||||
CAGTCCAGGCGCAAGTCCGCGGCACTGCTCCAGATGCTGCTGCTGCTGCTGCTGCTGCT 307
OY      687  |||||
CAGACGAGACCGGCGCAAGTCCGCGGCACTGCTCCAGATGCTGCTGCTGCTGCTGCTGCT 746
Db      308  |||||
CAGACGAGACCGGCGCAAGTCCGCGGCACTGCTCCAGATGCTGCTGCTGCTGCTGCTGCT 367
OY      747  |||||
CAACCGGTATACCAAGAGAGCTTCAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
Db      368  |||||
CAACCGGTATACCAAGAGAGCTTCAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
OY      807  |||||
TACCCAGAGCGCGGTACCCCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 866
Db      427  |||||
TACCCAGAGCGCGGTACCCCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
OY      867  |||||
TTTACAGATTCCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
Db      487  |||||
TTTACAGATTCCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
OY      927  |||||
TGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 986
Db      547  |||||
TGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
OY      987  |||||
CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1046
Db      607  |||||
CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
OY      1047  |||||
CATTAATTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1106
Db      667  |||||
CATTAATTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
OY      1107  |||||
GGCAATGAGGTTCTTTCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
Db      727  |||||
GGCAATGAGGTTCTTTCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
OY      1167  |||||
CATCGGATTCCTCCCAACATGCGGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1226
Db      787  |||||
CATCGGATTCCTCCCAACATGCGGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
OY      1227  |||||
TTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1261
Db      847  |||||
TTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 881

RESULT 4
BM922065 1044 bp mRNA linear EST 12-MAR-2002
LOCUS     BM922065
DEFINITION
AGENCOURT 6707015 NIH_MGC_115 Homo sapiens cDNA IMAGE:5754122
5', mRNA sequence.
ACCESSION BM922065
VERSION    BM922065.1 GI:19372444
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1044)
NIH-MGC http://mgs.nci.nih.gov/.

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12791 row: a column: 03
High quality sequence stop: 694.
Location/Qualifiers

FEATURES
source

```
1..1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5754122"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      23.3%; Score 741; DB 3; Length 1044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

202 AGTCGCTTCTGTCTGTGAGGGGGGCTGTGAGATTTCTAGTTATGATAATGAGACTTTA 261
23 AGGCTTCTGTCTGTGAGGGGGGCTGTGAGATTTCTAGTTATGATAATGAGACTTTA 82
262 AAAACGAGGAGCGGAGGCGAGGTTCAGGTTCTAGAGCTATGACGCTGAGACACTGCC 321
83 AAAACGAGGAGCGGAGGCGAGGTTCAGGTTCTAGAGCTATGACGCTGAGACACTGCC 142
322 TTTCTCTCTCTATCATGCTCTCCAGAAATTTCTCATATGTGACGACGACTACCAATT 381
143 TTTCTCTCTCTATCATGCTCTCCAGAAATTTCTCATATGTGACGACGACTACCAATT 202
382 CAGGGGGAGTCGAGGCTGTGTGACAGATCAATCCATTTCTCGACCACTGACAACTGG 441
203 CAGGGGGAGTCGAGGCTGTGTGACAGATCAATCCATTTCTCGACCACTGACAACTGG 262
442 AGAGAGATTACACTTTGAAAAAATTACAGAGGGGAGTGAAGATCAGTCAGATAGACACA 501
263 AGAGAGATTACACTTTGAAAAAATTACAGAGGGGAGTGAAGATCAGTCAGATAGACACA 322
502 ATTTTCTGACTCCAAAGACTACCAAGGGGAGCTCCAGAAAGTAACTCTCTCTGTT 561
323 ATTTTCTGACTCCAAAGACTACCAAGGGGAGCTCCAGAAAGTAACTCTCTCTGTT 382
562 TGGAGCGGGGTCTGTGAGCTTGTGACAGTTGATGGGCTGTGTGAGATGGCTACGCTC 621
383 TGGAGCGGGGTCTGTGAGCTTGTGACAGTTGATGGGCTGTGTGAGATGGCTACGCTC 442
622 TCTCTGAGTCAGCCAGCCACAGTCTGCGGCACTGCTCCAGTGCCATGTTCCCGTACC 681
443 TCTCTGAGTCAGCCAGCCACAGTCTGCGGCACTGCTCCAGTGCCATGTTCCCGTACC 502
682 CCGGCGAGCAAGGAGCGGCGACCCCGCTTCTTCATGCGGACGCTTACCGCTACATGG 741
503 CCGGCGAGCAAGGAGCGGCGACCCCGCTTCTTCATGCGGACGCTTACCGCTACATGG 562
```

```
QY 742 CCCACACCGGGTATCATCAACAGGAGCTTACAAACAGCTCTGTCCACTCTCTCGCGC 801
DB 563 CCCACACCGGGTATCATCAACAGGAGCTTACAAACAGCTCTGTCCACTCTCTCGCGC 622
QY 802 AGGATATCCCAACGCGCGGCTACCTCTTACCAACAGCATGAGGCACTCTTACCAAGAG 861
DB 623 AGGATATCCCAACGCGCGGCTACCTCTTACCAACAGCATGAGGCACTCTTACCAAGAG 682
QY 862 CTCGGTCTTACCACTTCCCTCCACGAGCGGGGCTGTGCGCGGCAAGACAGAGTGT 921
DB 683 CTCGGTCTTACCACTTCTCTCCACGAGCGGGGCTGTGCGCGGCAAGACAGAGTGT 742
QY 922 ACCTGTGCAACAGGCGCCCTTT 942
DB 743 ACCTGTGCAACAGGCGCCCTTT 763
```

```
RESULT 5
BX418703
LOCUS
DEFINITION BX418703 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
ACCESSION CS0DF0081F20 5-PRIME, mRNA sequence.
VERSION BX418703
KEYWORDS BX418703.2 GI:46957805
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1116)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30769503.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the PCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DF0081F20&cl=1851.f.
```

FEATURES
source

```
1..1116
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF0081F20"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the PCMVSPORT 6
vector. Library was not normalized."
```

ORIGIN

```
Query Match      22.8%; Score 725; DB 5; Length 1116;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 945; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

273 CCGGAGGCGAGGTTGAGTTCTAGAGCTATGAGAGCTGAGCACTGCTTCTCTCTCT 332
DB 9 CCGGAGGCGAGGTTGAGTTCTAGAGCTATGAGAGCTGAGCACTGCTTCTCTCTCT 68
QY 333 ATCATGCTCTTCAAGAAATTTCTCATATGTGAGAGCAAGCTTACCAATTCAGGCGGATCC 392
```

```

Db      69 ATCATGCTCTCCAGAAATTTTCTCATGTGAGACAGACGAGTACCCCAATTCAGGCGGATCC 128
QY      393 GAGCTTGTCTTTCACAGATCATTCATATCTTGACCACTGACACCTGGAGAAATTTCA 452
Db      129 GAGCTTGTCTTTCACAGATCATTCATATCTTGACCACTGACACCTGGAGAAATTTCA 188
QY      453 CTTTGGAAAAAATTAACAGGGGGATGACGATCATGTCAGATACAGAAATTTTCTTAC 512
Db      189 CTTTGGAAAAAATTAACAGGGGGATGACGATCATGTCAGATACAGAAATTTTCTTAC 248
QY      513 TCCAGGATCTACACAGGGGACGTCACAGAGATTAATCTCTCTGTTTGGACGGGGTC 572
Db      249 TCCAGGATCTACACAGGGGACGTCACAGAGATTAATCTCTCTGTTTGGACGGGGTC 308
QY      573 TCTGAGCTTCTGACAGATTCAGATGAGCTCTGTCGAGATGCTTACCTCTCTTCAATCC 632
Db      309 TCTGAGCTTCTGACAGATTCAGATGAGCTCTGTCGAGATGCTTACCTCTCTTCAATCC 368
QY      633 AGCCAGCCACAGTCTGCGGCGCACTGCTCCAGTGCATGTCCTCCGTAACCCGCGCAGAC 692
Db      369 AGCCAGCCACAGTCTGCGGCGCACTGCTCCAGTGCATGTCCTCCGTAACCCGCGCAGAC 428
QY      693 GGAACGGGACGACCCGCTTCTTCATCGGACGCTAGCCGCTACATGAGCCACACCGG 752
Db      429 GGAACGGGACGACCCGCTTCTTCATCGGACGCTAGCCGCTACATGAGCCACACCGG 488
QY      753 GTTCATCACCAAGAGCTTACCAAGAGCTCTGTCGACCTCTCTGCGGAGGATACCC 812
Db      489 GTTCATCACCAAGAGCTTACCAAGAGCTCTGTCGACCTCTCTGCGGAGGATACCC 547
QY      813 AGGCGCGCTACCCCTTACCAAGAGCTTACCAAGAGCTCTCTGCGGAGGATACCC 872
Db      548 AGGCGCGCTACCCCTTACCAAGAGCTTACCAAGAGCTCTCTGCGGAGGATACCC 607
QY      873 GAGTCTCTTCCACACGAGCGGCGCTGTCGCGGCGGACAGAGTACCTGTCGAC 932
Db      608 GAGTCTCTTCCACACGAGCGGCGCTGTCGCGGCGGACAGAGTACCTGTCGAC 667
QY      933 AGGCGCGCTTACCAAGAGCTTACCAAGAGCTTACCAAGAGCTTACCAAGAGG 992
Db      668 AGGCGCGCTTACCAAGAGCTTACCAAGAGCTTACCAAGAGCTTACCAAGAGG 727
QY      993 AGGCGATGTTCTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1052
Db      728 AGGCGATGTTCTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 787
QY      1053 AAATTTTGTGATGATGATTTTGGGATCCCAATCACTGAGGTTTCAAGAGGCA 1112
Db      788 AAATTTTGTGATGATGATTTTGGGATCCCAATCACTGAGGTTTCAAGAGGCA 847
QY      1113 TGGGTTCTTGGCGGCAAGCGACACCAATGTCAGAGAAATCGGGTCTATATGATCCG 1172
Db      848 TGGGTTCTTGGCGGCAAGCGACACCAATGTCAGAGAAATCGGGTCTATATGATCCG 907
QY      1173 GATTTCCCAACACTGGGGCTCACTGATGCGCAAGAAATCTCTTTT 1220
Db      908 GATTTCCCAACACTGGGGCTCACTGATGCGCAAGAAATCTCTTTT 955

```

```

RESULT 6
BX418656 824 bp mRNA linear EST 03-MAY-2004
LOCUS BX418656 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DP008YG10 5-PRIME, mRNA sequence.
ACCESSION BX418656
VERSION BX418656
KEYWORDS BX418656.2 GI:46954137
SOURCE EST.
ORGANISM Homo sapiens (human)
            Homo sapiens
            Rikaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo
REFERENCE 1 (bases 1 to 824)

```

```

AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30765911.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdnas=CS0DP008YD050P1&c=1851.f.
Location/Qualifiers
1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DP008YG10"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 22.0%; Score 701; DB 5; Length 824;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 821; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 203 GTGCTTCTGCTAGTAGAGGGGCTGTCGATTTCTTATGATTAATAGACTTTAA 262
Db 4 GTGCTTCTGCTAGTAGAGGGGCTGTCGATTTCTTATGATTAATAGACTTTAA 63
QY 263 AAACAGGAGCGGAGCGGAGGCTTCAAGTTCTAGACTTATGACCTGAGCACTGCT 322
Db 64 AAACAGGAGCGGAGCGGAGGCTTCAAGTTCTAGACTTATGACCTGAGCACTGCT 123
QY 323 TTCTCTCTTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGACGACTCCCAATTC 382
Db 124 TTCTCTCTTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGACGACTCCCAATTC 183
QY 383 AGGCGATCCGAGCTGTCCTTGGACAGATCATCCATTAATCTGACCACTGACCACTGGA 442
Db 184 AGGCGATCCGAGCTGTCCTTGGACAGATCATCCATTAATCTGACCACTGACCACTGGA 243
QY 443 GAGAGTTCACTTTGAAAAAATTAACAGGGGATGACGAATCATGTCAGATACAGACAA 502
Db 244 GAGAGTTCACTTTGAAAAAATTAACAGGGGATGACGAATCATGTCAGATACAGACAA 303
QY 503 TTTTCTGACTCAAGAGACTCAACGAGGAGCTCCAGAGAAATTAATCTCTCCGTCTT 562
Db 304 TTTTCTGACTCAAGAGACTCAACGAGGAGCTCCAGAGAAATTAATCTCTCTGTCTT 363
QY 563 GGAACGGGCTCTGAGCTTCTGACAGTTTGCATGAGCTCTGTCGAGATCGCTACTCTCT 622
Db 364 GGAACGGGCTCTGAGCTTCTGACAGTTTGCATGAGCTCTGTCGAGATCGCTACTCTCT 423
QY 623 CTCTGATCCAGCCAGCCACAGTCTGCGGCACTGCTCTCCAGTGCATGTCCTGATCC 682
Db 424 CTCTGATCCAGCCAGCCACAGTCTGCGGCACTGCTCTCCAGTGCATGTCCTGATCC 483
QY 683 CGGCGACAGCGACCGGCGGACCCGCTCTTCTCATCGGACGCTTACGCGCTTACATG 742
Db 484 CGGCGACAGCGACCGGCGGACCCGCTCTTCTCATCGGACGCTTACGCGCTTACATG 543
QY 743 CCAACGACCGGATCATCACCAAGAGGACTTAACAGCTCTGTCCTCTGCGCGCA 802

```

DB 544 CCACCAACCCGGTCAATCACCACGAGCCTACAAAGCCTCCTGCTCAACTCCTCGCCCA 603

QY 803 GGGATACCCCAAGCGCGGCTACCCCTACCCACAGAGTAAGGCACTCTTACCAAGAGC 862

DB 604 -GGATACCCCAAGCGCGGCTACCCCTACCCACAGAGTAAGGCACTCTTACCAAGAGC 662

QY 863 TCCGTTTCAACGAGTTCTCTTCCACACCGCGGGGCTGGTGGCCCGCAAGAGCAAGTGA 922

DB 663 TCCGTTTCAACGAGTTCTCTTCCACACCGCGGGGCTGGTGGCCCGCAAGAGCAAGTGA 722

QY 923 CCTGTGCAACAGAGCCCTTTGGCTGAATTTCAACGCGACCAAGAGAGATGATGAC 982

DB 723 CCTGTGCAACAGAGCCCTTTGGCTGAATTTCAACGCGACCAAGAGAGATGATGAC 782

QY 983 CAACAGGAGAGCGGCGATGTTCTTTTAAAGTTTAAAT 1024

DB 783 CAACAGGAGAGCGGCGATGTTCTTTTAAAGTTTAAAT 824

RESULT 7
DR001136 768 bp mRNA linear EST 17-MAY-2005
LOCUS TC115971 Human fetal brain, large insert, pCMV expression library
DEFINITION Homo sapiens cDNA clone TC115971 5' similar to Homo sapiens T-box, brain, 1 (TBR1), mRNA sequence.
ACCESSION DR001136
VERSION DR001136.1 GI:66261009
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 768)
AUTHORS Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L., Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M., Zhang,X., Jay,G. and He,W.
TITLE High-throughput cloning of full-length human cDNAs directly from cDNA libraries optimized for large and rare transcripts
JOURNAL Unpublished (2005)
COMMENT Contact: Kovacs, KF
High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: CDNA@origene.com
This EST submission is part of an on-going human full-length cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 5prime forward vector primer, Origene Technologies Inc.
FEATURES
source
1..768
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC115971"
/tissue_type="Fetal Brain"
/clone_lib="Human fetal brain, large insert, pCMV expression library"
/note="Organ: Fetal Brain; Vector: pCMV6-XL4; Site 1: EORI; Site 2: XhoI/SalI compatible end ligation; Oligo-dT primed reverse transcription optimized for large and GC rich mRNA transcripts, cDNA size selection, optimized ligation for large inserts into mammalian expression vector, random clones selected for end sequence

ORIGIN
Query Match 21.8%; Score 694; DB 8; Length 768;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

203 GTGCTTTCTGTCATGTCAGAGGGGCTGTGATTTCTGATTATGATAAGACTTTAA 262
9 GTGCTTTCTGTCATGTCAGAGGGGCTGTGATTTCTGATTATGATAAGACTTTAA 68

QY 263 AAACAGGAGAGGAGGCGAGTTCAGGTTCTAGAGCTATGAGAGCTGAGCACTGCT 322

DB 69 AAACAGGAGAGGAGGCGAGTTCAGGTTCTAGAGCTATGAGAGCTGAGCACTGCT 128

QY 323 TTTCCTTTCTATCATGCTCTTCCAGAAATTTCTCAATGTAGAGAGCACTACCAATTC 382

DB 129 TTTCCTTTCTATCATGCTCTTCCAGAAATTTCTCAATGTAGAGAGCACTACCAATTC 188

QY 383 AGGCGATCCGAGCTTGTCTTGACAGATCATCCATTATCTGACCACTGAACAACCTGGA 442

DB 189 AGGCGATCCGAGCTTGTCTTGACAGATCATCCATTATCTGACCACTGAACAACCTGGA 248

QY 443 GAGAAGTTCACTTTGAAAAAATTACAGGGGGATGACGAATCATGTCAGATACAGACAA 502

DB 249 GAGAAGTTCACTTTGAAAAAATTACAGGGGGATGACGAATCATGTCAGATACAGACAA 308

QY 503 TTTTCTTGATCTCAAGACTACAGAGGAGAGTCCAGAGAAAGTAACTTCTCTGCTT 562

DB 309 TTTTCTTGATCTCAAGACTACAGAGGAGAGTCCAGAGAAAGTAACTTCTCTGCTT 368

QY 563 GAGCGGGGTCCTGAGCTTGTGTCAGATTCAGTGGCTGTGTCAGATGCTGCTCT 622

DB 369 GAGCGGGGTCCTGAGCTTGTGTCAGATTCAGTGGCTGTGTCAGATGCTGCTCTCT 428

QY 623 CTCTCAGTCCAGCCAGCCAGCAGTCTGCGGCACCTCTCCAGTGCATGTTCCGTAACC 682

DB 429 CTCTCAGTCCAGCCAGCCAGCAGTCTGCGGCACCTCTCCAGTGCATGTTCCGTAACC 488

QY 683 CGGCGAGCAGGAGCCGCGGACCCCGCTTCCATGCGGACCTTACGCGCTACATGCG 742

DB 489 CGGCGAGCAGGAGCCGCGGACCCCGCTTCCATGCGGAGCCCTTACGCGCTACATGCG 548

QY 743 CCACCAACCCGGTCAATCAACAGGAGCTCAACAGGCTCTGTCACACTCTCGCGCA 802

DB 549 CCACCAACCCGGTCAATCAACAGGAGCTCAACAGGCTCTGTCACACTCTCGCGCA 608

QY 803 GGGATACCCCAAGCGCGGCTACCCCTACCCACAGAGTAAGGCACTCTTACCAAGAGC 862

DB 609 GGGATACCCCAAGCGCGGCTACCCCTACCCACAGAGTAAGGCACTCTTACCAAGAGC 668

QY 863 TCCGTTTCAACGAGTTCTCTTCCACACCGCGGGG 896

DB 669 TCCGTTTCAACGAGTTCTCTTCCACACCGCGGGG 702

RESULT 8
BX420367 921 bp mRNA linear EST 01-MAY-2004
LOCUS BX420367
DEFINITION CSODP027YH06 5-PRIME, mRNA sequence.
ACCESSION BX420367
VERSION BX420367.2 GI:46932405
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 921)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)

COMMENT

On May 13, 2003 this sequence version replaced gi:30650954.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 1851.f
 For more information about this cluster see
 http://www.genoscope.cns.fr/cdna?cs=CS0DF027D03QPI&c=1851.f.

FEATURES

source

1..921
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF027YH06"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo (dt) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN

Query Match 21.7%; Score 691; DB 5; Length 921;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 861; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

358 ATGTGAGGAGGAGCTACCAATTCAGGAGGATCCAGCTTGTCTTGACAGATATCCCA 417
 1 ATGTGAGGAGGAGCTACCAATTCAGGAGGATCCAGCTTGTCTTGACAGATATCCCA 60
 418 TTATCTCGACCACTGACCACTGAGAGAGTTCACCTTTGAAAAAATTCACAGGGAGA 477
 61 TTATCTCGACCACTGACCACTGAGAGAGTTCACCTTTGAAAAAATTCACAGGGAGA 120
 478 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537
 121 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 538 AGAGAGTAACT 597
 181 AGAGAGTAACT 240
 598 GCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 657
 241 GCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 658 CTCCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
 301 CTCCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 718 TCGGAGCTTACCGCTACATGAGCCACACCGGCTATACCAAGAGAGCTTACAACA 777
 361 TCGGAGCTTACCGCTACATGAGCCACACCGGCTATACCAAGAGAGCTTACAACA 420
 778 GCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
 421 GCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
 838 AGTACGGGCACTCTTACCAAGAGCTGTTTCAACAGTTCTCTTCAACCCAGCCGGGG 897
 480 AGTACGGGCACTCTTACCAAGAGCTGTTTCAACAGTTCTCTTCAACCCAGCCGGGG 539
 898 TGTGTCCTGGGCAAGGACAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 957
 540 TGTGTCCTGGGCAAGGACAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 599

QY 958 GGACCAAAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
 DB 600 GGACCAAAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
 QY 1018 TTACATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1077
 DB 660 TTACATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 719
 QY 1078 CGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1137
 DB 720 CGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779
 QY 1138 CCAATGTCAGAAAGATCGGCTATATGATCGGATTCGATTCGATTCGATTCGATTCGATTC 1197
 DB 780 CCAATGTCAGAAAGATCGGCTATATGATCGGATTCGATTCGATTCGATTCGATTCGATTC 839
 QY 1198 GGATGCGGCAAGAAATCTCTTTT 1220
 DB 840 GGATGCGGCAAGAAATCTCTTTT 862

RESULT 9
 B1914132 945 bp mRNA linear EST 16-OCT-2001
 LOCUS 603182077P1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246298 5',
 DEFINITION mRNA sequence.
 ACCESSION B1914132 GI:16178444
 VERSION B1914132.1 GI:16178444
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1 (bases 1 to 945)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LNL1621 row: 1 column: 19
 High quality sequence stop: 711.
 Location/Qualifiers

FEATURES

source

1..945
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246298"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_121"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH_MGC Library."

ORIGIN

Query Match 19.6%; Score 625; DB 3; Length 945;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 675; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 294 TCTAGAGCTATGACGTGGAGCAGTCCCTTCTCTCTATCATGCTCTCCAGAAATTT 353
DB 14 TCTTAGAGCTATGACGTGGAGCAGTCCCTTCTCTCTATCATGCTCTCCAGAAATTT 73
QY 354 CTCAATGTGAGCAGCAGTACCCACATTCAGCGGATCCGAGCTTGTCTTGACATCAT 413
DB 74 CTCAATGTGAGCAGCAGTACCCACATTCAGCGGATCCGAGCTTGTCTTGACATCAT 133
QY 414 CCCATTATTCGACCACTGACCAACCTTGAGAGAAATTTCACTTTGAAAAAATTTCCAGG 473
DB 134 CCCATTATTCGACCACTGACCAACCTTGAGAGAAATTTCACTTTGAAAAAATTTCCAGG 193
QY 474 GGGATGACGAATCATGTCAGATACAGCAATTTTCTGACTCCAGAGACTCACAGGAGAC 533
DB 194 GGGATGACGAATCATGTCAGATACAGCAATTTTCTGACTCCAGAGACTCACAGGAGAC 253
QY 534 GTCCAGAGAAGTAAACTCTCTCTGTCTGTGACGCGGCTCTGAGCTTGTGACAGTTTC 593
DB 254 GTCCAGAGAAGTAAACTCTCTCTGTCTGTGACGCGGCTCTGAGCTTGTGACAGTTTC 313
QY 594 GATGCTCTGCTGACAGATGCTACCTCTCTCTCTGACGTCAGCCAGCCAGCTGCGGCC 653
DB 314 GATGCTCTGCTGACAGATGCTACCTCTCTCTCTGACGTCAGCCAGCCAGCTGCGGCC 373
QY 654 ACTGCTCCAGTGCATGTTCCGTAACCCGCGCAGACGAGCCGCGACCCCGCTTC 713
DB 374 ACTGCTCCAGTGCATGTTCCGTAACCCGCGCAGACGAGCCGCGACCCCGCTTC 433
QY 714 TCCATCGGAGCCCTAGCCGCTACATGAGCCACACCCGCTCATACCAAGAGAGCTTAC 773
DB 434 TCCATCGGAGCCCTAGCCGCTACATGAGCCACACCCGCTCATACCAAGAGAGCTTAC 493
QY 774 AACAGCTCTCTGCTGACAGTCTCTGCGCGAGAGATACCCAGCGCGGCTGACCCCTACCA 833
DB 494 AACAGCTCTCTGCTGACAGTCTCTGCGCGAGAGATACCCAGCGCGGCTGACCCCTACCA 553
QY 834 CAGCAGTACGCGCAGTCTCTACCAAGAGTCCGTTCTTACAGTTCTCTCTCACCCAGCG 893
DB 554 CAGCAGTACGCGCAGTCTCTACCAAGAGTCCGTTCTTACAGTTCTCTCTCACCCAGCG 613
QY 894 GGGCTGTGCTCCGCGCAAGACAGTGTACTGTGTGACAGGCGCTTGTGCTGAAATTT 953
DB 614 GGGCTGTGCTCCGCGCAAGACAGTGTACTGTGTGACAGGCGCTTGTGCTGAAATTT 673
QY 954 CACCGGACCAAAACG 969
DB 674 CACCGGACCAAAACG 689

RESULT 10
BMS48197 1028 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6571595 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5733882
DEFINITION 5', mRNA sequence.
ACCESSION BMS48197
VERSION BMS48197.1 GI:18782590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1028)
NIH-MGC <http://mgi.mci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12738 row: e column: 19
High quality sequence stop: 623.
Location/Qualifiers
1..1028
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5733882"
/issue_type="hippocampus"
/lab_host="DH108"
/clone_id="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-Sport6; Site 1: EcorV
(destroyed); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

ORIGIN
Query Match 19.2%; Score 611; DB 3; Length 1028;
Best Local Similarity 100.0%; Pred. No. 1.3e-310;
Matches 611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 451 CACCTTTGAAAAAATTTACAGGGGGATGACGATTCGATACAGCAATTTCTCTG 510
DB 79 CACCTTTGAAAAAATTTACAGGGGGATGACGATTCGATACAGCAATTTCTCTG 138
QY 511 ACTCCAAGAGACTCACAGGGGACGTCCAGAGAAAGTAATCTCTCTGTCTTGACGGGG 570
DB 139 ACTCCAAGAGACTCACAGGGGACGTCCAGAGAAAGTAATCTCTCTGTCTTGACGGGG 198
QY 571 TCTGTAGCTTGTGACAGTTTGATGAGCTCTGTGTGACAGATCGTACTCTCTCACT 630
DB 199 TCTGTAGCTTGTGACAGTTTGATGAGCTCTGTGTGACAGATCGTACTCTCTCACT 258
QY 631 CCAGCCAGCCACATCTGCGGCGCATGCTCCAGATGCAATGTCGCGGCGGAGC 690
DB 259 CCAGCCAGCCACATCTGCGGCGCATGCTCCAGATGCAATGTCGCGGCGGAGC 318
QY 691 ACGGACGGGSCACCCGCTTCTTCATCGGACGCTTACCGGCTACATGAGCCACCAACC 750
DB 319 ACGGACGGGSCACCCGCTTCTTCATCGGACGCTTACCGGCTACATGAGCCACCAACC 378
QY 751 CGGTCAATCACCAGAGAGCTTACCAAGAGCTCTGTGCAATCTCTGCGGAGGATACC 810
DB 379 CGGTCAATCACCAGAGAGCTTACCAAGAGCTCTGTGCAATCTCTGCGGAGGATACC 438
QY 811 CCAGGCGCGGTACCCCTTACCCACAGAGTACGCGGCTCTCTACAGAGAGCTCCGTTCT 870
DB 439 CCAGGCGCGGTACCCCTTACCCACAGAGTACGCGGCTCTCTACAGAGAGCTCCGTTCT 498
QY 871 ACCAGTTCTCTCCACCAAGCGGGGCTGTGCGCGGCAAGACAGATGTACTGTGCA 930
DB 499 ACCAGTTCTCTCCACCAAGCGGGGCTGTGCGCGGCAAGACAGATGTACTGTGCA 558
QY 931 ACAGGCCCTTTGGCTGAATTTACCGGCAACCAAGAGATGATCATCAACAAACAG 990
DB 559 ACAGGCCCTTTGGCTGAATTTACCGGCAACCAAGAGATGATCATCAACAAACAG 618
QY 991 GAAGCGGATGTTCTTTTAAAGTTTAAATTTCTGTGCTCGATCCACGGCTCAT 1050
DB 619 GAAGCGGATGTTCTTTTAAAGTTTAAATTTCTGTGCTCGATCCACGGCTCAT 678
QY 1051 ACAATATTTT 1061
DB 679 ACAATATTTT 689

RESULT 11

BI914330 872 bp mRNA linear EST 17-OCT-2001
 LOCUS 603182311F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246210.5
 DEFINITION mRNA sequence.
 ACCESSION BI914330
 KEYWORDS EST.
 VERSION BI914330.1 GI:16198838
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 872)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgsabp-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11621 row: f column: 03
 High quality sequence stop: 853.
 Location/Qualifiers
 1..872
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246210"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_121"
 /notes="Organ: brain; Vector: pCMV-Sport6; Site_1: NciI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH_MGC Library."
 ORIGIN
 Query Match 18.5%; Score 588; DB 3; Length 872;
 Best Local Similarity 100.0%; Pred. No. 1.8e-298;
 Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 473 GGGGATGAGCAATCAGTACAGATACAGCAATTTTCTGACTCAAGAGACTCACCAGGGA 532
 105 GGGGATGAGCAATCAGTACAGATACAGCAATTTTCTGACTCAAGAGACTCACCAGGGA 164
 533 CGTCCAGAGAAATTAATCTCTCTCTGTTGACGGGGTCTGAGCTTGTCAAGTTT 592
 165 CGTCCAGAGAAATTAATCTCTCTCTGTTGACGGGGTCTGAGCTTGTCAAGTTT 224
 593 CGATGCTCTGTCGACATGCTTCTCTCTCACTCAAGCCAGCAACAGTCTGGCG 652
 225 CGATGCTCTGTCGACATGCTTCTCTCTCACTCAAGCCAGCAACAGTCTGGCG 284
 653 CACTGCTCCAGTGCATGTTCCCGTACCCCGGACAGACCGGCGACCCCGCTT 712
 285 CACTGCTCCAGTGCATGTTCCCGTACCCCGGACAGACCGGCGACCCCGCTT 344
 713 CTCATGAGGAGCCCTTAGCCGCTACATGAGCCACCAACCGGCTATCAACGAGCTTA 772
 345 CTCATGAGGAGCCCTTAGCCGCTACATGAGCCACCAACCGGCTATCAACGAGCTTA 404
 773 CAACAGCTCTCTGCTCAATCTCTGCGGAGGATACCCACGCGCGCTAACCCCTTAACC 832

Db 405 CAACAGCTCTCTGCTCAATCTCTGCGGACAGGATACCCACGCGCGCTAACCCCTTAACC 464
 833 ACAGAGTACGGCCATCTCTTACCAAGAGCTCCGTTCTTACAGTTCCTTCCACCCAGCC 892
 465 ACAGAGTACGGCCATCTCTTACCAAGAGCTCCGTTCTTACAGTTCCTTCCACCCAGCC 524
 893 GGGGCTGTCGCCGCAAGACAGAGTGTACTGTGCAACAGGCCCTTTGGCTGAATTT 952
 525 GGGGCTGTCGCCGCAAGACAGAGTGTACTGTGCAACAGGCCCTTTGGCTGAATTT 584
 953 TCACCGGACCAACCGAGATGATCATCAACCAACAGGAAAGCGCATGTTTCTTTT 1012
 585 TCACCGGACCAACCGAGATGATCATCAACCAACAGGAAAGCGCATGTTTCTTTT 644
 1013 AAGTTTAACTTTCTGCTCTGATCCACGCGCTATTAGCAATATTTT 1060
 645 AAGTTTAACTTTCTGCTCTGATCCACGCGCTATTAGCAATATTTT 692
 RESULT 12
 BP311133
 LOCUS BP311133 Sugano cDNA library, brain Homo sapiens cDNA clone
 DEFINITION NRR04714, mRNA sequence.
 ACCESSION BP311133
 VERSION BP311133.1 GI:52240108
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 583)
 Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 15342556
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp.
 Location/Qualifiers
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NRR04714"
 /tissue_type="brain"
 /clone_lib="Sugano cDNA library, brain"
 ORIGIN
 Query Match 18.2%; Score 581; DB 3; Length 583;
 Best Local Similarity 100.0%; Pred. No. 9.1e-295;
 Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 190 TTGCTGTTGAAGTCTTCTGCTAGTGAAGGGGTCTGGATTTCTAGTTATGATA 249
 3 TTGCTGTTGAAGTCTTCTGCTAGTGAAGGGGTCTGGATTTCTAGTTATGATA 62
 250 AATGAGCTTTAAAAACAGGAGCGGAGGGGAGTGTTCAGGTTCTAGAGCTTAGAGCC 309
 63 AATGAGCTTTAAAAACAGGAGCGGAGGGGAGTGTTCAGGTTCTAGAGCTTAGAGCC 122
 310 TGAAGACGCTTTCTCTTATCATGCTCTCAAGAAATTTCTCAATGAGACAGA 369
 123 TGAAGACGCTTTCTCTTATCATGCTCTCAAGAAATTTCTCAATGAGACAGA 182
 370 GCTACCAATTCAGGCGGATCGAGCTTGTCTTCAAGATCATCCATTATCTGACCA 429
 183 GCTACCAATTCAGGCGGATCGAGCTTGTCTTCAAGATCATCCATTATCTGACCA 242

QY 430 CTGACACCTGGAGAGAACTTCACTTTGAAAAAATTACAGGGGGATGACGATCGT 489
Db 243 CTGACAACTGGAGAGAACTTCACTTTGAAAAAATTACAGGGGGATGACGATCGT 302
QY 490 CAGATACAGACAAATTTCTGACTCAAGAGACTCAACGAGGAGCGTCCAGAACTAAC 549
Db 303 CAGATACAGACAAATTTCTGACTCAAGAGACTCAACGAGGAGCGTCCAGAACTAAC 362
QY 550 TCTCTCTGCTTTGGAAGGGGCTCTGAGCTTGTCAAGTTTCAGTGGCTCTGCTGAG 609
Db 363 TCTCTCTGCTTTGGAAGGGGCTCTGAGCTTGTCAAGTTTCAGTGGCTCTGCTGAG 422
QY 610 ATGCGTACTGCTCTCTGAGTCCAGAGCAAGTCCGAGGCACTGCTCCAGTTGCA 669
Db 423 ATGCGTACTGCTCTCTGAGTCCAGAGCAAGTCCGAGGCACTGCTCCAGTTGCA 482
QY 670 TGTTCCTGACCCCGGCGCAGCAGCAGCGGCGCACCCGCTTTCATCGGACGCTTA 729
Db 483 TGTTCCTGACCCCGGCGCAGCAGCAGCGGCGCACCCGCTTTCATCGGACGCTTA 542
QY 730 GCCGCTACATGGGCCACCAACCGGTCTATCAACAAGGAGCC 770
Db 543 GCCGCTACATGGGCCACCAACCGGTCTATCAACAAGGAGCC 583

RESULT 13
B1819019 639 bp mRNA linear EST 04-OCT-2001
LOCUS 603033141F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174290 5',
DEFINITION mRNA sequence.
ACCESSION B1819019
VERSION B1819019.1 GI:15930569
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 639)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov.
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1434 row: a column: 11
High quality sequence stop: 639.
Location/Qualifiers

FEATURES
source 1..639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174290"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

1. ORIGIN

Query Match 16.8%; Score 536; DB 3; Length 639;
Best Local Similarity 100.0%; Pred. No. 5.6e-271;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 GGGGATACGATATCAGTACAGTACAGCAAAATTTCTGACTCCAGGAGTACACAGGGA 532
Db 104 GGGGATACGATATCAGTACAGTACAGCAAAATTTCTGACTCCAGGAGTACACAGGGA 163
QY 533 CGTCCAGAGAAATTAATCTCTCTGCTTTGAGAGGGGCTCTGAGCTTGTCAAGTTT 592
Db 164 CGTCCAGAGAAATTAATCTCTCTGCTTTGAGAGGGGCTCTGAGCTTGTCAAGTTT 223
QY 593 CGATGGCTCTGTCAGATGAGTACCTCTCTCTGTCAGTCCAGGAGCAAGTCTGCGG 652
Db 224 CGATGGCTCTGTCAGATGAGTACCTCTCTCTGTCAGTCCAGGAGCAAGTCTGCGG 283
QY 653 CACTGCTCCAGTCCATGTTCCCGTACCCCGGCGACAGCGAGACCGGCGACCCGCTT 712
Db 284 CACTGCTCCAGTCCATGTTCCCGTACCCCGGCGACAGCGAGACCGGCGACCCGCTT 343
QY 713 CTCATGGGCGAGCCCTTACCGCTACATGGGCCACACCCGCTATCAACAAGAGCTTA 772
Db 344 CTCATGGGCGAGCCCTTACCGCTACATGGGCCACACCCGCTATCAACAAGAGCTTA 403
QY 773 CAACAGCCTCTCTGCAACTCTGCGCGCAGGAGTACCCACGCGCGGCTTACCCCTTACC 832
Db 404 CAACAGCCTCTCTGCAACTCTCTGCGCGCAGGAGTACCCACGCGCGGCTTACCCCTTACC 463
QY 833 ACACAGTACCGGCGACTCTTACCAAGAGTCCGTTTACAGTTCTCTTCCACCGAGCC 892
Db 464 ACACAGTACCGGCGACTCTTACCAAGAGTCCGTTTACAGTTCTCTTCCACCGAGCC 523
QY 893 GGGGCTGTCGCCCGGCAAGGACAGGTGTACTGTGCAAGGCGCCCTTTGGCTGAATT 952
Db 524 GGGGCTGTCGCCCGGCAAGGACAGGTGTACTGTGCAAGGCGCCCTTTGGCTGAATT 583
QY 953 TCACCGGCGACCAACGAGATGATCATCAACCAACAGGAGGCGGATTTCTT 1008
Db 584 TCACCGGCGACCAACGAGATGATCATCAACCAACAGGAGGCGGATTTCTT 639

RESULT 14
B1917781 740 bp mRNA linear EST 17-OCT-2001
LOCUS 603183881F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:524769 5',
DEFINITION mRNA sequence.
ACCESSION B1917781
VERSION B1917781.1 GI:16199720
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 740)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1625 row: b column: 22
High quality sequence stop: 740.
Location/Qualifiers

FEATURES
source 1..740
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5247669"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-Sport6; Site 1: Nct1;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 16.4%; Score 531; DB 3; Length 740;
Best Local Similarity 100.0%; Pred. No. 2.5e-268; Indels 0; Gaps 0;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 529 GGGAGCTCCAGAGAGTAACTCTCTCTGAGCGGGTCTCTGAGCTTCTGACA 588
DB 153 GGGAGCTCCAGAGAGTAACTCTCTCTGAGCGGGTCTCTGAGCTTCTGACA 212
QY 589 GTTTCATGCTCTGCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 648
DB 213 GTTTCATGCTCTGCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 272
QY 649 CGGCGCATGCTCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 708
DB 273 CGGCGCATGCTCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 332
QY 709 CTTTCATGCTCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 768
DB 333 CTTTCATGCTCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 392
QY 769 CTTTCATGCTCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 828
DB 393 CTTTCATGCTCTGAGATCGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 452
QY 829 ACCGACGAGTACGAGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 888
DB 453 ACCGACGAGTACGAGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 512
QY 889 AGCGGGGCTGAGTACGAGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 948
DB 513 AGCGGGGCTGAGTACGAGTACCTCTCTCTGAGCGGGTCTCTGAGCTT 572
QY 949 AATTTCACCGGCACTAAGGAGATGATCATCAACCAAGGAGGCGATGTTCTT 1008
DB 573 AATTTCACCGGCACTAAGGAGATGATCATCAACCAAGGAGGCGATGTTCTT 632
QY 1009 TTTTAAGTTTAACTTTGCTGATCCAGGCTCATTAATAATTT 1059
DB 633 TTTTAAGTTTAACTTTGCTGATCCAGGCTCATTAATAATTT 683

RESULT 15
AM003023/c 831 bp mRNA linear EST 08-MAR-2000
LOCUS
DEFINITION
w61d04.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2475751.3
similar to SW:TR1 HUMAN Q16650 T-BRAIN-1 PROTEIN; contains PIR5.13
TR1 repetitive element; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AM003023.1 GI:5849861
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 831)

AUTHORS
TITLE
JOURNAL
COMMENT
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLN at:
www.bio.lnlnl.gov/bbrp/image/image.html
Insert Length: 880 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 458.
Location/Qualifiers
1. 831

FEATURES

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2475751"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_1ib="NCI-CGAP GC6"
/note="Vector: pT73D-Pac (pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
as circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clone IDs
1257096-1258631, 1459064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 16.4%; Score 523; DB 1; Length 831;
Best Local Similarity 99.8%; Pred. No. 4.2e-264; Indels 0; Gaps 0;
Matches 573; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1364 CGCGGCGAGAGTACCTTCCCTGAGACTAGTTCATCGCGCTACCGCTACAGAA 1423
DB 616 CGCGGCGAGAGTACCTTCCCTGAGACTAGTTCATCGCGCTACCGCTACAGAA 557
QY 1424 CACGATATTACACAACTGAAATAGATCAACCTTTTGCAAAAGATTTCGGATTA 1483
DB 556 CACGATATTACACAACTGAAATAGATCAACCTTTTGCAAAAGATTTCGGATTA 497
QY 1484 TTATGACAGATCTACACCGGCTGTGACATGACCGCTGACCCCTCGCCCAAGATC 1543
DB 496 TTATGACAGATCTACACCGGCTGTGACATGACCGCTGACCCCTCGCCCAAGATC 437
QY 1544 GCCGCGCTGAGATGTCGCCGCGGCGCGCTACAGCATGCGCGCTTTCTTCGAGA 1603
DB 436 GCCGCGCTGAGATGTCGCCGCGGCGCGCTACAGCATGCGCGCTTTCTTCGAGA 377
QY 1604 CCAAGTTCGTAGCACTACGCAAGGCGCGCTTCAACCGCGCGCGCGCGCGCGCG 1663
DB 376 CCAAGTTCGTAGCACTACGCAAGGCGCGCTTCAACCGCGCGCGCGCGCGCGCG 317
QY 1664 GCCGCGTACGAGCGGAGCGTGTGCGCAACCAAGGCGCTGTGCGCGCAAGAGCCGA 1723
DB 316 GCCGCGTACGAGCGGAGCGTGTGCGCAACCAAGGCGCTGTGCGCGCAAGAGCCGA 257
QY 1724 GGAACCGCGCGCGCGCGCTGCGCGCAAGCGGTTTGTGACGCGCGCGCAACAGCGCTGGA 1783
DB 256 GGAACCGCGCGCGCGCGCTGCGCGCAAGCGGTTTGTGACGCGCGCGCAACAGCGCTGGA 197
QY 1784 CTTGCGGCGCTCGGCTATGACACGCGCAAGGATTCGCGGCGCAACGCGCGCAAGCTGCT 1843
DB 196 CTTGCGGCGCTCGGCTATGACACGCGCAAGGATTCGCGGCGCAACGCGCGCAAGCTGCT 137

Oy 1844 CTCTTACGCGGCGGCGGCGTGAAGGCGCTGCCCTGCAGGCTGCAGGCTGCACTGAGCG 1903
Db 136 CTCTTACGCGGCGGCGGCGTGAAGGCGCTGCCCTGCAGGCTGCAGGCTGCACTGAGCG 77
Oy 1904 CCGGCTCGGCTACTACGCGGACCGCTCGGCTG 1937
Db 76 CCGGCTCGGCTACTACGCGGACCGCTCGGCTG 43

Search completed: December 21, 2005, 01:04:47
Job time : 12035 secs

Db 241 GACTCAGCAGGGAGCGTCCAGAGTAACCTCTCTCTGTGTGAGCGGGCTCTAG 300
OY 579 CTTTCGTACAGTTTTCGATGGCTCTGTGAGATTCGTACCTCTCTTCATGTCAGCCAG 638
Db 301 CTTTCGTACAGTTTTCGATGGCTCTGTGAGATTCGTACCTCTCTTCATGTCAGCCAG 360
OY 639 CCAAGTCGCGGCCCATCTGCTCCAGTGCATGTTCCCGTACCCGGCCAGACGAGACCG 698
Db 361 CCAAGTCGCGGCCCATCTGCTCCAGTGCATGTTCCCGTACCCGGCCAGACGAGACCG 420
OY 699 GCGACCCCGCCTTCTCATTCGAGAGCCCTAGCCGCTACATGAGCCACCCGGTCAATC 758
Db 421 GCGACCCCGCCTTCTCATTCGAGAGCCCTAGCCGCTACATGAGCCACCCGGTCAATC 480
OY 759 ACCAAGGAGCCTACACAGCCTCTGTCCAACTCTCGCGGAGGATATCCCAAGGCG 818
Db 481 ACCAAGGAGCCTACACAGCCTCTGTCCAACTCTCGCGGAGGATATCCCAAGGCG 540
OY 819 GCGTACCCCTACCCAGAGATACGCGCACTCTACCAAGAGCTCCGTTCAACGATTC 878
Db 541 GCGTACCCCTACCCAGAGATACGCGCACTCTACCAAGAGCTCCGTTCAACGATTC 600
OY 879 TCCTCACCACGAGCGGGGCTGTGCGCGGCAAGCAGGTATCCTGTGCAACAGGCGC 938
Db 601 TCCTCACCACGAGCGGGGCTGTGCGCGGCAAGCAGGTATCCTGTGCAACAGGCGC 660
OY 939 CTTTGGCTGAAATTTTACCGGCGACCAACGAGATGATCATCAACAAAGGAGGCG 998
Db 661 CTTTGGCTGAAATTTTACCGGCGACCAACGAGATGATCATCAACAAAGGAGGCG 720
OY 999 ATGTTCCCTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTT 1058
Db 721 ATGTTCCCTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTT 780
OY 1059 TTTGTGATGATGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGGTT 1118
Db 781 TTTGTGATGATGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGGTT 840
OY 1119 CCTTGGCGGCAAGCGGACCAATGTGCAAG---GAAATCGGGTCTATATGATCCGGA 1174
Db 841 CCTTGGCGGCAAGCGGACCAATGTGCAAG---GAAATCGGGTCTATATGATCCGGA 900
OY 1175 TTTCCCGCAACACTGGGGGCTCACTGATGCGCAAGAAATCTCTTTTGGAAAATTAAACT 1234
Db 901 TTTCCCGCAACACTGGGGGCTCACTGATGCGCAAGAAATCTCTTTTGGAAAATTAAACT 960
OY 1235 TAGAACAACAAGAGCTTCAATTAACAATGAGGAGATGATGTTTCAAGTCTTGA 1294
Db 961 TAGAACAACAAGAGCTTCAATTAACAATGAGGAGATGATGTTTCAAGTCTTGA 1020
OY 1295 CAAATACAGCCCGCTGATGTGTGAAGTGAACGAGACGCGACGAGAGACATAG 1354
Db 1021 CAAATACAGCCCGCTGATGTGTGAAGTGAACGAGAGCGGACGAGAGACATAG 1080
OY 1355 CCAAGCCCGCGGTCAGACCTTCACTTTCCCTGAGACTCAAGTTATGCGGTACCGC 1414
Db 1081 CCAAGCCCGCGGTCAGACCTTCACTTTCCCTGAGACTCAAGTTATGCGGTACCGC 1140
OY 1415 CTACAGAAACAGGATATTACAACTGAAAATAGATCAAACTTTTGGAAAAGATT 1474
Db 1141 CTACAGAAACAGGATATTACAACTGAAAATAGATCAAACTTTTGGAAAAGATT 1200
OY 1475 TCGGATTAATTATGACAGATCTTACACGAGCTGTGACATGACCGGCTGACCCCTCGC 1534
Db 1201 TCGGATTAATTATGACAGATCTTACACGAGCTGTGACATGACCGGCTGACCCCTCGC 1260
OY 1535 CAAAGACTCGCGCGCTGTGAGATGTGTGCGCGGGCGGCTAGCCATGCGCGCTCTT 1594
Db 1261 CAAAGACTCGCGCGCTGTGAGATGTGTGCGCGGGCGGCTAGCCATGCGCGCTCTT 1320
OY 1595 CCGCAGAGACAGTGTGAGCAACTACGCAAGGCGCGCTTCAACCGGCGCGGGCGCG 1654
Db 1321 CCGCAGAGACAGTGTGAGCAACTACGCAAGGCGCGCTTCAACCGGCGCGGGCGCG 1380

OY 1655 GGGCCCCGGGCGGCTACGACCGGACGTCGCGCACACCAAGGGGCTGTCGCGCA 1714
Db 1381 GGGCCCCGGGCGGCTACGACCGGACGTCGCGCACACCAAGGGGCTGTCGCGCA 1440
OY 1715 GCAAGCGAGAGACCGGGGCGGCGCTGCGGCAACGCTGTTTGTGAAGCGCGCAACAA 1774
Db 1441 GCAAGCGAGAGACCGGGGCGGCGCTGCGGCAACGCTGTTTGTGAAGCGCGCAACAA 1500
OY 1775 CCGGCTGAGACTTGGCGGCTTGGGCTTATGACAGGCAAGACTTTCGCGGCAACGCGC 1834
Db 1501 CCGGCTGAGACTTGGCGGCTTGGGCTTATGACAGGCAAGACTTTCGCGGCAACGCGC 1560
OY 1835 CAGCGTGTCTTAAACGGGCGGCGGAGTGAAGCGCTGCGCTGAGGCTGAGGCTG 1894
Db 1561 CAGCGTGTCTTAAACGGGCGGCGGAGTGAAGCGCTGCGCTGAGGCTGAGGCTG 1620
OY 1895 CACTGAGCGCGCGCTGAGCTTACAGCGGACCCGTCGAGGCTGAGGCGCGCACTCCCG 1954
Db 1621 CACTGAGCGCGCGCTGAGCTTACAGCGGACCCGTCGAGGCTGAGGCGCGCACTCCCG 1680
OY 1955 GCAAGTCTGCGGCAACAAATGCGGCTGAGTGCCTGCTGCGCAACAGCGCGCGC 2014
Db 1681 GCAAGTCTGCGGCAACAAATGCGGCTGAGTGCCTGCTGCGCAACAGCGCGCGC 1740
OY 2015 CGCGCGCGGCAATGCGCGCGCAATCCCTTACCTGGGCGAGAGAGCGAGGCGCTGCGC 2074
Db 1741 CGCGCGCGGCAATGCGCGCGCAATCCCTTACCTGGGCGAGAGAGCGAGGCGCTGCGC 1800
OY 2075 CGAGGCTGCGCGGCTGCGCGCGCGCGCGCGAGAGAGCGCAAGGCGCAAGGACTGCGA 2134
Db 1801 CGAGGCTGCGCGGCTGCGCGCGCGCGCGCGAGAGAGCGCAAGGCGCAAGGACTGCGA 1860
OY 2135 TTTCACTGATGATGAGAGCGCTTCTTGATCAATGATCAATGATCAATGATCAATGAT 2194
Db 1861 TTTCACTGATGATGAGAGCGCTTCTTGATCAATGATCAATGATCAATGATCAATGAT 1920
OY 2195 TTTAGAGAGGCGCAAGCGGAGAGTCTGCGCGGCGGCAAGCGCGGTCGAGAGTTC 2254
Db 1921 TTTAGAGAGGCGCAAGCGGAGAGTCTGCGCGGCGGCAAGCGCGGTCGAGAGTTC 1980
OY 2255 GTCCCGCTCAAGAGCGAGGTGTGCGCGCAAGCGGAGCTGCGAGAGAACTGCGCAAGA 2314
Db 1981 GTCCCGCTCAAGAGCGAGGTGTGCGCGCGCGCGCGAGAGAACTGCGCAAGA 2040
OY 2315 CATTAAGCGGCTACTATGCTTCTACTGTGCAAGTGAAGCGCGCGCTGCGCGCGCGC 2374
Db 2041 CATTAAGCGGCTACTATGCTTCTACTGTGCAAGTGAAGCGCGCGCTGCGCGCGCGC 2100
OY 2375 GCGCGGCGCGGACCCCGAGCAGCCCTCAACAGTCTTCCGAGCTCCGCTCCGACA 2434
Db 2101 GCGCGGCGCGGACCCCGAGCAGCCCTCAACAGTCTTCCGAGCTCCGCTCCGACA 2160
OY 2435 CTCCTCTTGGGCAACCACTCAATTTATTTAGCCCTGATGCGGCTGCGAGCAATAG 2494
Db 2161 CTCCTCTTGGGCAACCACTCAATTTATTTAGCCCTGATGCGGCTGCGAGCAATAG 2220
OY 2495 TGCAGTCTCCGAGCGGATTTTAACTTTTTCAGACAGAGTCTGCAATTAGCTCAC 2554
Db 2221 TGCAGTCTCCGAGCGGATTTTAACTTTTTCAGACAGAGTCTGCAATTAGCTCAC 2280
OY 2555 CGACTTCAACTTTGCTGTAAACCTTTTGTGTTTCTACTTAACCTTCTGTGAGATT 2614
Db 2281 CGACTTCAACTTTGCTGTAAACCTTTTGTGTTTCTACTTAACCTTCTGTGAGATT 2340
OY 2615 ATTCCTTAACAATTCCTTCCCTCGCTTCTTCTTAACCTTCTTAACCTTCTTCTTGTGA 2674
Db 2341 ATTCCTTAACAATTCCTTCCCTCGCTTCTTCTTAACCTTCTTCTTCTTCTTGTGA 2400
OY 2675 ATGAACCTTTCACCTTTAGAGACCTGGGAGTCTGTGAGGACAGGCAATTCGAGC 2734
Db 2401 ATGAACCTTTCACCTTTAGAGACCTGGGAGTCTGTGAGGACAGGCAATTCGAGC 2460

QY 2735 GCGCAAGTCTGGGCTCCACATTAAACATAGATGTTGACTGTAGAACCTGAGACCCACC 2794
 DB 2461 CCCCAAGTCTGGGCTCCACATTAAACATAGATGTTGACTGTAGAACCTGAGACCCACC 2520
 QY 2795 AACGGTCTCTTCTTATCCCGAGTGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2854
 DB 2521 AACGGTCTCTTCTTATCCCGAGTGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2580
 QY 2855 AATTTAGTGAACAAAGCCTGTGAATGTTGATGATGTTGATGATGTTGATGATGTTGATGAT 2914
 DB 2581 AATTTAGTGAACAAAGCCTGTGAATGTTGATGATGTTGATGATGTTGATGATGTTGATGAT 2640
 QY 2915 TGGCTAGTGTATTTATCTGCTCAAGGCAAAACCACTGATGCTTAACTTTTTCCTT 2974
 DB 2641 TGGCTAGTGTATTTATCTGCTCAAGGCAAAACCACTGATGCTTAACTTTTTCCTT 2700
 QY 2975 TCCCTTCTTGGCTTCTTCTCTCTCTCATACCTTCTCTCTCTCTTAACTTTTC 3034
 DB 2701 TCCCTTCTTGGCTTCTTCTCTCTCTCATACCTTCTCTCTCTCTTAACTTTTC 2760
 QY 3035 TTGTGATATATATTTCTAAGAGGCTCTAAGAAATGAAATGATGATGATGATGATGATGAT 3094
 DB 2761 TTGTGATATATTTCTAAGAGGCTCTAAGAAATGAAATGATGATGATGATGATGATGAT 2820
 QY 3095 CCACTTCTCTCAATCCGTTGATGAATATATTTACTATGCTTAACTGACACAAATG 3154
 DB 2821 CCACTTCTCTCAATCCGTTGATGAATATATTTACTATGCTTAACTGACACAAATG 2880
 QY 3155 CTAAAGAGATCCACCCAAACCTTTAAAG 3186
 DB 2881 CTAAAGAGATCCACCCAAACCTTTAAAG 2912

RESULT 2

US-09-949-016-16576
 ; Sequence 16576, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16576
 ; LENGTH: 12681
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-16576

Query Match 53.2%; Score 1694; DB 3; Length 12681;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1493 GATCTACACCGGCTGTGATGACGCGCTGACCCCTCGCCCAACGACTCGCGGCTC 1552
 DB 8988 GATCTACACCGGCTGTGATGACGCGCTGACCCCTCGCCCAACGACTCGCGGCTC 9047
 QY 1553 GCAGATGTCGCCGCGGCGCTGACGCGCTGACGCGCTGTCCTTCTGACGACGATTCGT 1612
 DB 9048 GCAGATGTCGCCGCGGCGCTGACGCGCTGACGCGCTGTCCTTCTGACGACGATTCGT 9107
 QY 1613 GAGGAACTTACGACGCGGCTGTCGACGCGGCGCGGCGCGGCGCGGCGCGGCGGCTAC 1672
 DB 9108 GAGGAACTTACGACGCGGCTGTCGACGCGGCGCGGCGCGGCGCGGCGCGGCGGCTAC 9167

QY 1673 GGAACCGAGCGTGGCGGACACCAACGAGGCTGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1732
 DB 9168 GGAACCGAGCGTGGCGGACACCAACGAGGCTGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 9227
 QY 1733 GCGGCTCTGCGCGGACGCTGTTTGTGACGCGCGGACCAACGCGCTGACCTTTCGCGGC 1792
 DB 9228 GCGGCTCTGCGCGGACGCTGTTTGTGACGCGCGGACCAACGCGCTGACCTTTCGCGGC 9287
 QY 1793 CTCGCGCTTATGACACGCGGACGCTTGTGCGGCGGACGCGGCGGACGCTGCTTCTTATGCG 1852
 DB 9288 CTCGCGCTTATGACACGCGGACGCTTGTGCGGCGGACGCGGCGGACGCTGCTTCTTATGCG 9347
 QY 1853 GCGGCGGCGGCTGAAAGGCGCTGCGGCGGACGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1912
 DB 9348 GCGGCGGCGGCTGAAAGGCGCTGCGGCGGACGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9407
 QY 1913 CTACTACGCGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1972
 DB 9408 CTACTACGCGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9467
 QY 1973 GTGCGGCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2032
 DB 9468 GTGCGGCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9527
 QY 2033 GCGCAATCCCTTACCTGCGGCGGAGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2092
 DB 9528 GCGCAATCCCTTACCTGCGGCGGAGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9587
 QY 2093 GCGCGGCGGCGGCGGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 2152
 DB 9588 GCGCGGCGGCGGCGGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 9647
 QY 2153 GCCCTCTGATCAAGTCAATGACTGCAAGGCTGCGGAGTTTACGAGCGAGGCGCAAGCG 2212
 DB 9648 GCCCTCTGATCAAGTCAATGACTGCAAGGCTGCGGAGTTTACGAGCGAGGCGCAAGCG 9707
 QY 2213 GAGGCGGATCTGCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2272
 DB 9708 GAGGCGGATCTGCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9767
 QY 2273 GGTGCTGGCGGCGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2332
 DB 9768 GGTGCTGGCGGCGGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9827
 QY 2333 CTTCTACTGCAAGCTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2392
 DB 9828 CTTCTACTGCAAGCTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9887
 QY 2393 AGCGAGCGGCTGACAGCTCTTCCCGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2452
 DB 9888 AGCGAGCGGCTGACAGCTCTTCCCGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 9947
 QY 2453 CTCATTTTATTTGACCGCTGATGCGGCTGCGGAGGAAATGATGAGGCTTCCGAGCGTG 2512
 DB 9948 CTCATTTTATTTGACCGCTGATGCGGCTGCGGAGGAAATGATGAGGCTTCCGAGCGTG 10007
 QY 2513 AATTTAACCTTTTGTGACAGGAGCTCTGCAATTAGCTACCGACTTCAACTTTGCTG 2572
 DB 10008 AATTTAACCTTTTGTGACAGGAGCTCTGCAATTAGCTACCGACTTCAACTTTGCTG 10067
 QY 2573 TAAACCTTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2632
 DB 10068 TAAACCTTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 10127
 QY 2633 TCCCGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2692
 DB 10128 TCCCGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 10187
 QY 2693 AGGAGACTGAGGAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2752
 DB 10188 AGGAGACTGAGGAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10247

```

Db      274  CGAAGCGTGGTTTGTGAGCGCCGSCCAACACGGCTGAGACTTCGCGGCTCGACTATG 215
QY      1804  ACAAGGSCACAGGACTTTCGCGGSCCAACGCGGSCACGCTGCTCTTACGCGGCGCGCGCG 186
Db      214  ACAAGGSCACAGGACTTTCGCGGSCCAACGCGGSCACGCTGCTCTTACGCGGCGCGCGCG 155
QY      1864  TGAAGGCGGTGGCGGCTGACAGAGGTGACAGGCTACATGGCGCGCCGCGTGGCTACTAGAGCG 192
Db      154  TGAAGGCGGTGGCGGCTGACAGAGGTGACAGGCTACATGGCGCGCCGCGTGGCTACTAGAGCG 95
QY      1924  ACCCGTCGGGCTTGCGGCGCGCGCGCACTGCCCGCCAGTACTGCGGACCCAACTCGGCGTCCG 198
Db      94  ACCCGTCGGGCTTGCGGCGCGCGCGCACTGCCCGCCAGTACTGCGGACCCAACTCGGCGTCCG 35
QY      1984  TGCTGCGCTTGCTGCGCCCAACAGCGCGCGGCGCGCG 2017
Db      34  TGCTGCGCTTGCTGCGCCCAACAGCGCGCGGCGCGCG 1

RESULT 4
US-09-513-999C-885
; Sequence 885, Application US/0951399C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513.999C

```

```

; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 885
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 117..386
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4
; OTHER INFORMATION: w=a or t
US-09-513-999C-885

```

	Matches	368;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
OY	191	TTGCTGTTGAAGTGCCTTTCTGTCTAGTGAAGGGGCTCTGGATTTCTAGTTTATGATTA	250							
Db	5	TTGCTGTTGAAGTGCCTTTCTGTCTAGTGAAGGGGCTCTGGATTTCTAGTTTATGATTA	64							
OY	251	ATGAGACTTAAAAACGAGGACGGGAGGGGAGTGTTCAGTTTCTAGACTATGAGAGCT	310							
Db	65	ATGAGACTTAAAAACGAGGACGGGAGGGGAGTGTTCAGTTTCTAGACTATGAGAGCT	124							
OY	311	GGAGCACTCCCTTTCTCTCTCTATATATCTCTCCAGAAATTTCTCAATGTGACGACAG	370							
Db	125	GGAGCACTCCCTTTCTCTCTCTATATATCTCTCCAGAAATTTCTCAATGTGACGACAG	184							
OY	371	CTACCCACATTCAGGCGGATTCGAGCTGTCTTGCACGATTCATCCATTATCTCGACAC	430							
Db	185	CTACCCACATTCAGGCGGATTCGAGCTGTCTTGCACGATTCATCCATTATCTCGACAC	244							
OY	431	TGACAACTTGGAGAGAGAGTTCACTTTTGAATAAATTAACAAGGGGATGACGAATCAATC	490							
Db	245	TGACAACTTGGAGAGAGAGTTCACTTTTGAATAAATAACAAGGGGATGACGAATCAATC	304							


```

/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13722
; LENGTH: 31469
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13722

Query Match 0.8%; Score 27; DB 3; Length 31469;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2842
DB 19253 GAGTGATGATGATGATGATG 19279

RESULT 9
US-09-949-016-13646/c
; Sequence 13646, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13646
; LENGTH: 23640
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(23640)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13646

Query Match 0.8%; Score 26; DB 3; Length 23640;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
DB 7919 GAGTGATGATGATGATGATG 7894

RESULT 10
US-09-949-016-16389
; Sequence 16389, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16389
; LENGTH: 26314
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16389

Query Match 0.8%; Score 26; DB 3; Length 26314;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
DB 5513 GAGTGATGATGATGATGATG 5538

RESULT 11
US-09-949-016-13006
; Sequence 13006, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13006
; LENGTH: 41243
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13006

Query Match 0.8%; Score 26; DB 3; Length 41243;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
DB 150 GAGTGATGATGATGATGATG 175

RESULT 12
US-09-949-016-13424/c
; Sequence 13424, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13424
; LENGTH: 85850

```
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(85850)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13424

Query Match          0.8%; Score 26; DB 3; Length 85850;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2816 GAGTGGATGGATGGATGGATGGATGG 2841
DB      5916 GAGTGGATGGATGGATGGATGGATGG 5891

RESULT 13
US-09-949-016-12362
; Sequence 12362, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12362
; LENGTH: 85912
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12362

Query Match          0.8%; Score 26; DB 3; Length 85912;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2819 TGGATGGATGGATGGATGGATGGATGG 2844
DB      20265 TGGATGGATGGATGGATGGATGGATGG 20290

RESULT 14
US-09-949-016-16109
; Sequence 16109, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16109
; LENGTH: 85913
; TYPE: DNA
```

```
; ORGANISM: Human
US-09-949-016-16109

Query Match          0.8%; Score 26; DB 3; Length 85913;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2819 TGGATGGATGGATGGATGGATGGATGG 2844
DB      20265 TGGATGGATGGATGGATGGATGGATGG 20290

RESULT 15
US-09-949-016-150203/C
; Sequence 150203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150203
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150203

Query Match          0.8%; Score 25; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2819 TGGATGGATGGATGGATGGATGGATGG 2843
DB      185 TGGATGGATGGATGGATGGATGGATGG 161
```

Search completed: December 21, 2005, 01:13:54
Job time : 545 secs

This Page Blank (uspto)

Db 61 CAGGAGCCGTTTCATCGGCTGCACAAGCAGCAAGATCAAAAGTGAGCCTTTTCTGATTGC 120
Qy 121 TGCATAGTGTCAATTGGCCCAATCTCTTCCCAAGGAAAAAAGTAAATCAACCTT 180
Db 121 TGCATAGTGTCAATTGGCCCAATCTCTTCCCAAGGAAAAAAGTAAATCAACCTT 180
Qy 181 TGAGAGCATTGTGCTGTGAAGTCTTCTGTCTAGTAGAGGGGTCTGTGATTTCTAG 240
Db 181 TGAGAGCATTGTGCTGTGAAGTCTTCTGTCTAGTAGAGGGGTCTGTGATTTCTAG 240
Qy 241 TTTTATGATTAATAGACTTTTAAAAACAGAGGACGGGAGCGAGTTCAGGTTCTAG 300
Db 241 TTTTATGATTAATAGACTTTTAAAAACAGAGGACGGGAGCGAGTTCAGGTTCTAG 300
Qy 301 CTATGAGGTGAGGACGTGCTTCTCTCTATCATGCTCTCCAAAGAAATTTCTCATAG 360
Db 301 CTATGAGGTGAGGACGTGCTTCTCTCTATCATGCTCTCCAAAGAAATTTCTCATAG 360
Qy 361 TGAGCAGCAGTACCCACATTCAGGCGGATCCGAGCTTGTCTTGGACGATCCCATTTA 420
Db 361 TGAGCAGCAGTACCCACATTCAGGCGGATCCGAGCTTGTCTTGGACGATCCCATTTA 420
Qy 421 TCTCGACCATGACAACTCTGAGAGAGTTCACCTTTGAAAAAAATTAACAGGGGAGTGA 480
Db 421 TCTCGACCATGACAACTCTGAGAGAGTTCACCTTTGAAAAAAATTAACAGGGGAGTGA 480
Qy 481 CGAATCAGTCAGATCAGACAAATTTTCTGACTCCAAAGGACTCACAGGGGACGTCCAGA 540
Db 481 CGAATCAGTCAGATCAGACAAATTTTCTGACTCCAAAGGACTCACAGGGGACGTCCAGA 540
Qy 541 GAAAGTAACTCTCTCTGTCTTGAACGGGGTCTGAGGCTTGTGACAGTTCGATGAGT 600
Db 541 GAAAGTAACTCTCTCTGTCTTGAACGGGGTCTGAGGCTTGTGACAGTTCGATGAGT 600
Qy 601 CTGCTGCGAGATGCTAATCTCTCTCTCAATCCAGCCAGCACAAGTGTGGGCCACTGCTC 660
Db 601 CTGCTGCGAGATGCTAATCTCTCTCTCAATCCAGCCAGCACAAGTGTGGGCCACTGCTC 660
Qy 661 CGAGTGCACATGTTCCCGTACCCCGGACGACGAGACCGGCGACCCCGCTTCTCATAG 720
Db 661 CGAGTGCACATGTTCCCGTACCCCGGACGAGACCGGCGACCCCGCTTCTCATAG 720
Qy 721 GCAGCCCTAGCCGCTACATGAGCCACACCCGGTCAATCAACAGAGCTTACAAAGCC 780
Db 721 GCAGCCCTAGCCGCTACATGAGCCACACCCGGTCAATCAACAGAGCTTACAAAGCC 780
Qy 781 TCTGTTCACATCTCTGCGCGCAGAGATACCCACGGCGGGCTTACCCCTTACCAAGCAT 840
Db 781 TCTGTTCACATCTCTGCGCGCAGAGATACCCACGGCGGGCTTACCCCTTACCAAGCAT 840
Qy 841 ACGGGCACATCTACCAAGGAGCTCGGTTCTACAGTTCCTTCAACCCAGCGCGGAGCTGG 900
Db 841 ACGGGCACATCTACCAAGGAGCTCGGTTCTACAGTTCCTTCAACCCAGCGCGGAGCTGG 900
Qy 901 TGCCCGGCAAGACAGGTGTACTGTGCAACAGGCCCTTGTGAGTGAATTTCAACCGAC 960
Db 901 TGCCCGGCAAGACAGGTGTACTGTGCAACAGGCCCTTGTGAGTGAATTTCAACCGAC 960
Qy 961 ACCAAACGAGATGATCATCACCAACAGGGAAGCGCATGTTTCTTTTAAAGTTTAA 1020
Db 961 ACCAAACGAGATGATCATCACCAACAGGGAAGCGCATGTTTCTTTTAAAGTTTAA 1020
Qy 1021 ACATTTGTGATGATGATCCAGGGCTATTAATATTTTGTGATGATGATTTGGGGG 1080
Db 1021 ACATTTGTGATGATGATCCAGGGCTATTAATATTTTGTGATGATGATTTGGGGG 1080
Qy 1081 ATCCCAATCATGAGGTTTCAAGAGGCAAAATGGGTTCTTGGCGGCAAGCGGACACA 1140
Db 1081 ATCCCAATCATGAGGTTTCAAGAGGCAAAATGGGTTCTTGGCGGCAAGCGGACACA 1140
Qy 1141 ATGTCCAAAGAAATGGGTCTATATGATCCGGAATTTCCCAACACTGGGGCTCACTGGA 1200
Db 1141 ATGTCCAAAGAAATGGGTCTATATGATCCGGAATTTCCCAACACTGGGGCTCACTGGA 1200

Qy 1201 TGCCCAAGAAATCTTTTGGAAAAATTAACCTTACGAACAAAGAGACTTCAATA 1260
Db 1201 TGCCCAAGAAATCTTTTGGAAAAATTAACCTTACGAACAAAGAGACTTCAATA 1260
Qy 1261 ACAATGGGACAGATGGTGTTTTACAGTCTTGGCAAAATACAGACCCCGCTGCATGTGG 1320
Db 1261 ACAATGGGACAGATGGTGTTTTACAGTCTTGGCAAAATACAGACCCCGCTGCATGTGG 1320
Qy 1321 TGGAAATGAAAGAGAGCGGACCGAGGACATTAACCGCCCGGCGGTGAGACGTTTCA 1380
Db 1321 TGGAAATGAAAGAGAGCGGACCGAGGACATTAACCGCCCGGCGGTGAGACGTTTCA 1380
Qy 1381 CTTTCCCTGAGACTCAGTTCATGCGCCGTCACCGGCTACAGAACAGGATTTTACACAC 1440
Db 1381 CTTTCCCTGAGACTCAGTTCATGCGCCGTCACCGGCTACAGAACAGGATTTTACACAC 1440
Qy 1441 TGAATAATGATCACACCCCTTTTGCAAAAGGATTTGGGATTAATATGACAGATCTTACA 1500
Db 1441 TGAATAATGATCACACCCCTTTTGCAAAAGGATTTGGGATTAATATGACAGATCTTACA 1500
Qy 1501 CCGGCTGTGACATGAGACCGCTGACCCCTTGCCCAAGATTCGCGCGGCTGCGAGATCG 1560
Db 1501 CCGGCTGTGACATGAGACCGCTGACCCCTTGCCCAAGATTCGCGCGGCTGCGAGATCG 1560
Qy 1561 TGCCCGGGGCGCGCTACGACATGAGCGGCTCTTCTCTGACAGACCAAGTTCGTCAGCACT 1620
Db 1561 TGCCCGGGGCGCGCTACGACATGAGCGGCTCTTCTCTGACAGACCAAGTTCGTCAGCACT 1620
Qy 1621 ACGCCAAGGCCGCTTCCACCACCGGGGCGGAGCGGAGCCCGGAGCGGATACGGAACGCA 1680
Db 1621 ACGCCAAGGCCGCTTCCACCACCGGGGCGGAGCGGAGCCCGGAGCGGATACGGAACGCA 1680
Qy 1681 GCGTGCAGCACAACCAACGGGCTGCTGTCGCGCAGACAGCCGAGAACCCGGGCGGCGCT 1740
Db 1681 GCGTGCAGCACAACCAACGGGCTGCTGTCGCGCAGACAGCCGAGAACCCGGGCGGCGCT 1740
Qy 1741 CGCCGCAACGCTGTGTTTGTGACGCGGCGCAACACCGGCTTGAATTGCGGCGCTCGAGCT 1800
Db 1741 CGCCGCAACGCTGTGTTTGTGACGCGGCGCAACACCGGCTTGAATTGCGGCGCTCGAGCT 1800
Qy 1801 ATGACACGCGCACGGAATTCGCGGGCAACGCGGACGCTGCTCTTTACGCGGCGGCGG 1860
Db 1801 ATGACACGCGCACGGAATTCGCGGGCAACGCGGACGCTGCTCTTTACGCGGCGGCGG 1860
Qy 1861 GCGTGAAGCGCTCCGCTGAGAGCTGAGGCTGCACTGGCGGCGGCTCGGCTACTACG 1920
Db 1861 GCGTGAAGCGCTCCGCTGAGAGCTGAGGCTGCACTGGCGGCGGCTCGGCTACTACG 1920
Qy 1921 CCGACCCGCTCGGGCTGGGGGCGCCGCAAGTCCCGCAGTACTGCGGACCAAGTCGAGGCT 1980
Db 1921 CCGACCCGCTCGGGCTGGGGGCGCCGCAAGTCCCGCAGTACTGCGGACCAAGTCGAGGCT 1980
Qy 1981 CGGTGTGCTCTGTGTGCGCCAAACGAGCGCGGCGCGCGCATGAGCGCGCCCAATC 2040
Db 1981 CGGTGTGCTCTGTGTGCGCCAAACGAGCGCGGCGCGCGCATGAGCGCGCCCAATC 2040
Qy 2041 CTTACTGTGGGGAAGAGCGGAGGCTGTGGCGCGGAGCGCTCGCCCTGCGCGCGGCGG 2100
Db 2041 CTTACTGTGGGGAAGAGCGGAGGCTGTGGCGCGGAGCGCTCGCCCTGCGCGCGGCGG 2100
Qy 2101 CCGCGGAGAGCGCAAGGCCCAAGGACTGTCCGATTCAGTTCAGTGTGATGAGAGCGGCTCT 2160
Db 2101 CCGCGGAGAGCGCAAGGCCCAAGGACTGTTCGATTCAGTTCAGTGTGATGAGAGCGGCTCT 2160
Qy 2161 CGATCAAGTCCATGATCCAGCGACTCGGGGATTTTACGAGAGCCCAAGGAGGCGGGA 2220
Db 2161 CGATCAAGTCCATGATCCAGCGACTCGGGGATTTTACGAGAGCCCAAGGAGGCGGGA 2220
Qy 2221 TCTGCGCGGCGGACACCGCCGCTGTCCGAGATTTGTTCCCGCTCAAGAGCGAGGTGCTGG 2280
Db 2221 TCTGCGCGGCGGACACCGCCGCTGTCCGAGATTTGTTCCCGCTCAAGAGCGAGGTGCTGG 2280

QY 2281 CCCACGGGACCTGCGAAGAACTGCGCCAGAGATTAGCGGCTTACTATGCTCTCT 2340
 Db 2281 CCCACGGGACCTGCGAAGAACTGCGCCAGAGATTAGCGGCTTACTATGCTCTCT 2340
 QY 2341 CGCAGAGTGGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2400
 Db 2341 CGCAGAGTGGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2400
 QY 2401 CCTCAGAGCTCTTCCCGAGCTCCGCTCCCGACACTCTCTCTTGGCGACCCACTATTT 2460
 Db 2401 CCTCAGAGCTCTTCCCGAGCTCCGCTCCCGACACTCTCTCTTGGCGACCCACTATTT 2460
 QY 2461 ATTGACCTCGATGGCGCTGCGACGGAATAGTCAAGTCTCCGAGGTGATTTTAC 2520
 Db 2461 ATTGACCTCGATGGCGCTGCGACGGAATAGTCAAGTCTCCGAGGTGATTTTAC 2520
 QY 2521 CTTTGTGACAGAGCTCTGCAATTTAGTCAACGAGCTTCACTTCTGTTAAACCTT 2580
 Db 2521 CTTTGTGACAGAGCTCTGCAATTTAGTCAACGAGCTTCACTTCTGTTAAACCTT 2580
 QY 2581 TTGGTTTCTTACTTACTCTTCTTCTGAGAGTTATCTCTCAATTTCCCTCCCTC 2640
 Db 2581 TTGGTTTCTTACTTACTCTTCTTCTGAGAGTTATCTCTCAATTTCCCTCCCTC 2640
 QY 2641 GTCTTCTCTTACTCTTACTCTTCTTCTGTAATGAATCTTCACTTTAGAGAGC 2700
 Db 2641 GTCTTCTCTTACTCTTACTCTTCTTCTGTAATGAATCTTCACTTTAGAGAGC 2700
 QY 2701 TGGGAGCTCTGTCAG 2760
 Db 2701 TGGGAGCTCTGTCAG 2760
 QY 2761 CATAGAGTGTGACTGTAGAACCTGAGCCACCCAGCGGCTCTTCTTATCCCGAGT 2820
 Db 2761 CATAGAGTGTGACTGTAGAACCTGAGCCACCCAGCGGCTCTTCTTATCCCGAGT 2820
 QY 2821 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
 Db 2821 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
 QY 2881 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
 Db 2881 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
 QY 2941 ACAAAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
 Db 2941 ACAAAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
 QY 3001 TCTCATCT 3060
 Db 3001 TCTCATCT 3060
 QY 3061 TAAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 3120
 Db 3061 TAAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 3120
 QY 3121 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
 Db 3121 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
 QY 3181 TAAAGG 3186
 Db 3181 TAAAGG 3186

RESULT 2
 US-10-029-386-20697
 ; Sequence 20697, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn. Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G;
 ; FILE REFERENCE: AROMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 20697
 ; LENGTH: 864
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC009487.3
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
 ; OTHER INFORMATION: NT HIT: g1142971, EVALUATE 0.00e+00
 ; OTHER INFORMATION: SWISSPROT HIT: Q1650, EVALUATE 0.00e+00
 ; OTHER INFORMATION: EST_HUMAN HIT: AI201654.1, EVALUATE 0.00e+00
 ; US-10-029-386-20697
 Query Match 27.1%; Score 864; DB 6; Length 864;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 864; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1493 GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCCAAAGACTGCGCGCTC 1552
 Db 1 GATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCCAAAGACTGCGCGCTC 60
 QY 1553 GCAAGTCTGCG 1612
 Db 61 GCAAGTCTGCG 120
 QY 1613 GAGCAACTACGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1672
 Db 121 GAGCAACTACGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 1673 GAGCGGAGCGTGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 1732
 Db 181 GAGCGGAGCGTGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 240
 QY 1733 CGCGCGCTGCG 1792
 Db 241 CGCGCGCTGCG 300
 QY 1793 CTGCGGCTATGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1852
 Db 301 CTGCGGCTATGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 1853 GCGCGGCGCGTGAAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1912
 Db 361 GCGCGGCGCGTGAAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 QY 1913 CTACTACGCGACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1972
 Db 421 CTACTACGCGACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 QY 1973 GTGCGGCTGCGGCTGCG 2032
 Db 481 GTGCGGCTGCGGCTGCG 540
 QY 2033 CGCAATCCCTTACTTGGGCGAGAGCGCGAGGCGCTGCGCGCGCGCGCGCGCGCGCGCG 2092
 Db 541 CGCAATCCCTTACTTGGGCGAGAGCGCGAGGCGCTGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 2093 GCCCGGCG 2152
 Db 601 GCCCGGCG 660
 QY 2153 GCCCTCTCGATCAAGTCCATGCACTCAGGCACTCGGAGATTAGAGAGGCGCAAGCG 2212

DB 661 GGCCTCTGATCAATGATCATGACTCCAGCAGCTCGGGATTTCAGAGAGCCAAAGC 720
QY 2213 GAGCGGATCTCGCGCGCGGACACCGCGTGTCCAGAGTTGTCGCCGCTCAAGAGCA 2272
DB 721 GAGCGGATCTCGCGCGCGGACACCGCGTGTCCAGAGTTGTCGCCGCTCAAGAGCA 780
QY 2273 GGTGTGCGCCGAGCGGAGTGCAGAGAACTGCGCAAGACATTAGCGGCTACTATGG 2332
DB 781 GGTGTGCGCCGAGCGGAGTGCAGAGAACTGCGCAAGACATTAGCGGCTACTATGG 840
QY 2333 CTTTACTCGCAGACTAGGCGCG 2356
DB 841 CTTTACTCGCAGACTAGGCGCG 864

RESULT 3
US-10-029-386-6986
; Sequence 6986, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; PRIORITY FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 6986
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009487.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EST HUMAN HIT: H10108.1, EVALU8 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q16650, EVALU8 5.00e-15
; OTHER INFORMATION: NT HIT: g111429971, EVALU8 0.00e+00
US-10-029-386-6986

Query Match 16.2%; Score 515; DB 6; Length 515;
Best Local Similarity 100.0%; Pred. No. 1,1e-257;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2236 CGCCCGTGTCCAGAGTTGTCGCCGCTCAAGAGCGAGGTGTGCGCCAGCGGAGCTGGC 2295
DB 1 CGCCCGTGTCCAGAGTTGTCGCCGCTCAAGAGCGAGGTGTGCGCCAGCGGAGCTGGC 60
QY 2296 AGAAGAACTGCGCGCAAGACATTAGCGGCTACTATGCTTCTTACTGCGACAGCTAGCGC 2355
DB 61 AGAAGAACTGCGCGCAAGACATTAGCGGCTACTATGCTTCTTACTGCGACAGCTAGCGC 120
QY 2356 CCCCTGCGCGCGCGCGCGCGCGCGCGCGAGACCCCGACGAGCCCGCTCAGAGCTTTCC 2415
DB 121 CCCCTGCGCGCGCGCGCGCGCGCGCGCGAGACCCCGACGAGCCCGCTCAGAGCTTTCC 180
QY 2416 CGAGCTCGCGCTCCCGACACTCTCTTTCGCGACCCGACTCATTTTATTTGACCTCGAG 2475
DB 181 CGAGCTCGCGCTCCCGACACTCTCTTTCGCGACCCGACTCATTTTATTTGACCTCGAG 240
QY 2476 GCGGCTGCGAGGAAATAGTGAAGTCTCCGAGCGTATTTTAACTTTTGGACAGCA 2535
DB 241 GCGGCTGCGAGGAAATAGTGAAGTCTCCGAGCGTATTTTAACTTTTGGACAGCA 300
QY 2536 GTCCTGCAATTAAGTCAACGACTTCACTTTGCTGAACCTTTTGGTTTCTACTT 2595

DB 301 GTCCTGGAATTAAGTCAACGACTTCACTTTGCTGAACCTTTTGGTTTCTACTT 360
QY 2596 ACTCTTCTTCTGTGAGATTAATCTCTCAACAATTCCTCCCTCGCTTTCTTACTT 2655
DB 361 ACTCTTCTTCTGTGAGATTAATCTCTCAACAATTCCTCCCTCGCTTTCTTACTT 420
QY 2656 CCTACTTCTCTTCTTGTGAATGAACCTTCACTTTAGAGAGAGCTGGAGTCTGTCA 2715
DB 421 CCTACTTCTCTTCTTGTGAATGAACCTTCACTTTAGAGAGAGCTGGAGTCTGTCA 480
QY 2716 GGCAGACGCAATTCGACCGCGCAAGTCTCGGCT 2750
DB 481 GGCAGACGCAATTCGACCGCGCAAGTCTCGGCT 515

RESULT 4
US-10-842-072-37/c
; Sequence 37, Application US/10842072
; Publication No. US20050158731A1
; GENERAL INFORMATION:
; APPLICANT: PLASS, CHRISTOPH
; TITLE OF INVENTION: DETECTION OF METHYLATED CpG RICH SEQUENCES DIAGNOSTIC FOR
; FILE REFERENCE: 22727-04242
; CURRENT APPLICATION NUMBER: US/10/842,072
; PRIORITY FILING DATE: 2004-05-10
; PRIOR APPLICATION NUMBER: 09/775,398
; PRIORITY FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 37
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-842-072-37

Query Match 16.1%; Score 514; DB 9; Length 514;
Best Local Similarity 100.0%; Pred. No. 3,6e-257;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1504 GCTGTGACATGAGACCGGCTGACCCCTCGCCCAAGACATGCGCGCTCGCAGATGCTGC 1563
DB 514 GCTGTGACATGAGACCGGCTGACCCCTCGCCCAAGACATGCGCGCTCGCAGATGCTGC 455
QY 1564 CCGGGGCGCGTACGACATGCGCGGCTTTCTCTGAGAGACCAATTTGTAGCAACTACG 1623
DB 454 CCGGGGCGCGTACGACATGCGCGGCTTTCTCTGAGAGACCAATTTGTAGCAACTACG 395
QY 1624 CCAAGGCGCGCTTCACCGCGGCGCGGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 1683
DB 394 CCAAGGCGCGCTTCACCGCGGCGCGGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 335
QY 1684 TGCAGCAACCAAGCGGCTGTCTGTGCGCGAGAGGCGAGGACCGGCGCGGCTTGC 1743
DB 334 TGCAGCAACCAAGCGGCTGTCTGTGCGCGAGAGGCGAGGACCGGCGCGGCTTGC 275
QY 1744 CGCAACGCTGTTGTGACGCGCGGCGCAACAACCGGCTGAGATTGGGAGCTTGGCTATG 1803
DB 274 CGCAACGCTGTTGTGACGCGCGGCGCAACAACCGGCTGAGATTGGGAGCTTGGCTATG 215
QY 1804 ACAGGCGGACGAGACTTGGCGGCGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1863
DB 214 ACAGGCGGACGAGACTTGGCGGCGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 155
QY 1864 TGAAGCGCTGCGCTGAGGCTGCAAGCTGACCTGCGCGGCGGCTGAGCTACTACGCG 1923
DB 154 TGAAGCGCTGCGCTGAGGCTGCAAGCTGACCTGCGCGGCGGCTGAGCTACTACGCG 95
QY 1924 ACCGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1983
DB 94 ACCGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 35


```

: APPLICANT: Alexander Olek
: APPLICANT: Christian Piepenbrock
: APPLICANT: Kurt Berlin
: TITLE OF INVENTION: Method for determining the degree of methylation of defined
: TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3'
: FILE REFERENCE: E01/1227
: CURRENT APPLICATION NUMBER: US/10/363,345A
: CURRENT FILING DATE: 2003-03-03
: NUMBER OF SEQ ID NOS: 40712
: SEQ ID NO 9002
: LENGTH: 1026
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
: OTHER INFORMATION: CpG-island No: 9002
US-10-363-345A-9002

```

Query Match 1.0%; Score 32; DB 8; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCCTGCGCCACGACTGCGCGCTCGCA 1555
|||||
DB 112 ACCCCCTGCGCCACGACTGCGCGCTCGCA 143

RESULT 9
US-10-363-483A-9001/c
; Sequence 9001, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9001
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 9001
US-10-363-483A-9001

Query Match 1.0%; Score 32; DB 9; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCCTGCGCCACGACTGCGCGCTCGCA 1555
|||||
DB 915 ACCCCCTGCGCCACGACTGCGCGCTCGCA 864

RESULT 10
US-10-363-483A-9002
; Sequence 9002, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9002
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 9002
US-10-363-483A-9002

Query Match 1.0%; Score 32; DB 9; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1524 ACCCCCTGCGCCACGACTGCGCGCTCGCA 1555
|||||
DB 112 ACCCCCTGCGCCACGACTGCGCGCTCGCA 143

RESULT 11
US-10-027-632-273001/c
; Sequence 273001, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 273001
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-273001

Query Match 0.8%; Score 27; DB 5; Length 1006;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2819 TGGATGATGATGATGATGATGATGATG 2845
|||||
DB 894 TGGATGATGATGATGATGATGATGATG 868

RESULT 12
US-10-027-632-273001/c
; Sequence 273001, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 273001
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-273001

Query Match 0.8%; Score 27; DB 6; Length 1006;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2819 TCGATGATGATGATGATGATGATG 2845
 ||||||||||||||||||
 DB 894 TCGATGATGATGATGATGATGATG 868

RESULT 13

US-10-674-124A-16640
 ; Sequence 16640, Application US/10674124A
 ; Publication No. US2004019797A1
 ; GENERAL INFORMATION:

APPLICANT: INOKO, Hidetoshi
 APPLICANT: TAMITA, Gen
 TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
 FILE REFERENCE: ORIN-003CIP

CURRENT APPLICATION NUMBER: US/10/674,124A
 CURRENT FILING DATE: 2003-09-26
 PRIOR APPLICATION NUMBER: 10/257,511
 PRIOR FILING DATE: 2003-03-07
 PRIOR APPLICATION NUMBER: PCT/JP00/07621
 PRIOR FILING DATE: 2000-10-30
 PRIOR APPLICATION NUMBER: JP2000-112699
 PRIOR FILING DATE: 2000-04-13
 PRIOR APPLICATION NUMBER: JP2002-327516
 PRIOR FILING DATE: 2002-09-28
 PRIOR APPLICATION NUMBER: JP2002-383869
 PRIOR FILING DATE: 2002-12-09
 NUMBER OF SEQ ID NOS: 27110
 SEQ ID NO 16640

LENGTH: 182
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: D20S208
 FEATURE:
 OTHER INFORMATION: Located on chromosome 10
 FEATURE:
 OTHER INFORMATION: Distance between a terminus base of telomere on
 OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
 OTHER INFORMATION: sequence : 103456393
 FEATURE:
 OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
 OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
 OTHER INFORMATION: 5'-terminus of this base sequence : 52125
 US-10-674-124A-16640

Query Match 0.8%; Score 26; DB 8; Length 182;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
 ||||||||||||||||||
 DB 107 GAGTGATGATGATGATGATG 132

RESULT 14

US-10-719-993-31636/c
 ; Sequence 31636, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 FILE REFERENCE: CL001496
 CURRENT APPLICATION NUMBER: US/10/719,993
 CURRENT FILING DATE: 2003-11-24
 NUMBER OF SEQ ID NOS: 55342

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 31636
 LENGTH: 201
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-719-993-31636

Query Match 0.8%; Score 26; DB 8; Length 201;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
 ||||||||||||||||||
 DB 150 GAGTGATGATGATGATGATG 125

RESULT 15

US-10-719-993-31637/c
 ; Sequence 31637, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 FILE REFERENCE: CL001496
 CURRENT APPLICATION NUMBER: US/10/719,993
 CURRENT FILING DATE: 2003-11-24
 NUMBER OF SEQ ID NOS: 55342
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 31637
 LENGTH: 201
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-719-993-31637

Query Match 0.8%; Score 26; DB 8; Length 201;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2816 GAGTGATGATGATGATGATG 2841
 ||||||||||||||||||
 DB 149 GAGTGATGATGATGATGATG 124

Search completed: December 21, 2005, 01:54:54
 Job time : 2454 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 07:04:33 ; Search time 15679 Seconds
(without alignments)
11550.688 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186
Sequence: 1 caggtgattacctaataa.....caccacaacaccttaagg 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sta.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2884.4	90.5	2910	6	CQ729238 Sequence
2	2753	86.4	2894	8	HSU49250 Human putat
3	2490	78.2	3992	9	BC052737 Mus muscu
4	2400.4	75.3	3897	9	BC058399 Mus muscu
5	2334.8	73.3	3814	9	MMU49251 Mus muscu
6	1793	56.3	1793	6	CQ769652 Sequence
7	1793	56.3	1793	6	BC029289 Homo sapi
8	1695.4	53.2	18366	8	AC009487 Homo sapi
9	1289	40.5	101040	14	AC165777 Bos tauri
10	1164.6	36.6	209994	9	AL845291 Mouse DNA
11	1157.8	36.3	176172	14	AC127930 Rattus no
12	971	30.5	101040	14	AC165777 Bos tauri
13	694.8	21.8	2218	5	AF287006 Danio rer
14	514	16.1	514	6	AR655897 Sequence
15	491.2	15.4	793	8	HS4341671 Homo sapi
16	420.4	13.2	2900	5	XUW5996 Xenopus lae
17	417.6	13.1	690	8	HS433162 Homo sapi
18	412	12.9	3038	5	AB019785 Cynops py

19	407.6	12.8	2986	5	BC084243 Xenopus 1
20	407.4	12.8	2480	5	AF329830 Danio rer
21	407.4	12.8	2572	5	BC067719 Danio rer
22	404.2	12.7	2450	5	AF287007 Danio rer
23	401.6	12.6	689	8	HS4324262 Homo sapi
24	395.8	12.4	2842	5	CT025431 Xenopus t
25	386.2	12.1	641	8	HS438878 Homo sapi
26	384.6	12.1	388	6	BD024632 Sequence
27	384.6	12.1	388	6	AX885022 Sequence
28	374.6	11.8	2494	6	AR169779 Sequence
29	374.6	11.8	2840	8	BC037568 Homo sapi
30	374.6	11.8	3577	9	BC094319 Mus muscu
31	356.6	11.2	784	8	HS4327080 Homo sapi
32	353	11.1	1894	6	CQ723583 Sequence
33	350.8	11.0	2883	8	AB031037 Mus muscu
34	346	10.9	2061	8	AB031038 Homo sapi
35	341.2	10.7	2756	6	CS035351 Sequence
36	341.2	10.7	2756	6	CS036706 Sequence
37	341.2	10.7	2756	6	CS044303 Sequence
38	341.2	10.7	2756	6	CS045658 Sequence
39	330.2	10.4	741	8	HS433254 Homo sapi
40	328.2	10.3	691	8	HS4338038 Homo sapi
41	326.2	10.2	579	9	AY457971 Rattus no
42	321.8	10.1	693	8	AY4341518 Homo sapi
43	321.4	10.1	540	9	AF013281 Mus muscu
44	316.2	9.9	1608	6	AX055554 Sequence
45	316.2	9.9	1608	6	AX695541 Sequence

ALIGNMENTS

RESULT 1
LOCUS CQ729238 2910 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15172 from Patent WO02068579.
ACCESSION CQ729238
VERSION CQ729238.1 GI:42300290

KEYWORDS

ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: WO 02068579-A 15172 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES

source

1..2910
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 90.5%; Score 2884.4; DB 6; Length 2910;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2907; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY	279	GCGAGTGTTCAGGTTCTGAGCTATGACGCTGAGACATCGCTTCTCTTATCAG	338
DB	1	GCGGAGTGTTCAGGTTCTGAGCTATGACGCTGAGACATCGCTTCTCTTATCAG	60
QY	339	CTCTCAAGAAATTTCTCATGAGGACAGACGCTACCCATTCAGCGGATCGAGCTT	398
DB	61	CTCTCAAGAAATTTCTCATGAGGACAGACGCTACCCATTCAGCGGATCGAGCTT	120
QY	399	GTCCTGACGATCATCCATTTATCTGACCACTGACCACTGAGAGAAAGTTACCTTGG	458
DB	121	GTCCTGACGATCATCCATTTATCTGACCACTGACCACTGAGAGAAAGTTACCTTGG	180

QY 459 AAAAAAAAAATTTACAGGGGGGATGAOGAATCATGATACAGACAAATTTTCTGACTCCAG 518
Db 181 AAAAAAAAAATTTACAGGGGGGATGAOGAATCATGATACAGACAAATTTTCTGACTCCAG 240
QY 519 GACTCAGCAGGGGAGCTCCAGAGAAATAACTCTCTCTGCTTGGACGGGCTCTGAG 578
Db 241 GACTCAGCAGGGGAGCTCCAGAGAAATAACTCTCTCTGCTTGGACGGGCTCTGAG 300
QY 579 CTTTGTGACAGTTTGGATGAGCTCTGTGTGAGATTCGTAACCTCTCTCTCACTGACCCAG 638
Db 301 CTTTGTGACAGTTTGGATGAGCTCTGTGTGAGATTCGTAACCTCTCTCTCACTGACCCAG 360
QY 639 CCACAGTCTGGGGCACTGCTCCAGTGCATGTTCCGGTACCCCGGGCAGACGGAGCCG 698
Db 361 CCACAGTCTGGGGCACTGCTCCAGTGCATGTTCCGGTACCCCGGGCAGACGGAGCCG 420
QY 699 GGGCAGCCCCGCTTCTCATTCGAGAGCCCTAGCCGGTACATGAGCCACACCCGGTCAATC 758
Db 421 GGGCAGCCCCGCTTCTCATTCGAGAGCCCTAGCCGGTACATGAGCCACACCCGGTCAATC 480
QY 759 ACCAAGCAGGCTTACACAGGCTCTGTCTCAACTCTGTGCGCAGGAGATACCCACGGCC 818
Db 481 ACCAAGCAGGCTTACACAGGCTCTGTCTCAACTCTGTGCGCAGGAGATACCCACGGCC 540
QY 819 GGGTACCCCCCTACCCAGAGATAGGGCACTCTCTACCAAGAGCTCCGTTCTACAGATTG 878
Db 541 GGGTACCCCCCTACCCAGAGATAGGGCACTCTCTACCAAGAGCTCCGTTCTACAGATTG 600
QY 879 TCCTCCACCAGCCGGGGCTGTGTGCTCCGCAAGCAAGTGTACCTGTGCAACAGGCC 938
Db 601 TCCTCCACCAGCCGGGGCTGTGTGCTCCGCAAGCAAGTGTACCTGTGCAACAGGCC 660
QY 939 CTTTGGCTGAATTTTCACTGGGCAACCAAGAGATGATATATCAACAAAGGAAAGGGCC 998
Db 661 CTTTGGCTGAATTTTCACTGGGCAACCAAGAGATGATATATCAACAAAGGAAAGGGCC 720
QY 999 ATGTTTCCCTTTTAACTTTTAACTTTCTGATCTCGATCCAGGGCTATTACAAATTT 1058
Db 721 ATGTTTCCCTTTTAACTTTTAACTTTCTGATCTCGATCCAGGGCTATTACAAATTT 780
QY 1059 TTTGTGATGATGATTTTGGCGGATCCCAATCACTGAGGGTTTCAAGAGGCAATGGGTT 1118
Db 781 TTTGTGATGATGATTTTGGCGGATCCCAATCACTGAGGGTTTCAAGAGGCAATGGGTT 840
QY 1119 CTTTGGCGGAAGCGGACCAATGTGCAAGAAATCGGCTATATATGATCCGATTC 1178
Db 841 CTTTGGCGGAAGCGGACCAATGTGCAAGAAATCGGCTATATATGATCCGATTC 900
QY 1179 CCCAACACTGGGGCTCACTGATGCGCCAAAG-AAATCTTTTGGAAATTTAAACTTAC 1237
Db 901 CCCAACACTGGGGCTCACTGATGCGCCAAAGAAATCTTTTGGAAATTTAAACTTAC 960
QY 1238 GAACAACAAA- GAGAGCTTCAAAATAACAATGAGCAAGTGTGTTTACAGTCTTGACA 1296
Db 961 GAACAACAAAAGGAGCTTCAAAATAACAATGAGCAAGTGTGTTTACAGTCTTGACA 1020
QY 1297 AGTACCAAGCCCGCTGATGTGTGTGAAGTGAAGAGACGGCACCGAGACATAGCC 1356
Db 1021 AGTACCAAGCCCGCTGATGTGTGTGAAGTGAAGAGACGGCACCGAGACATAGCC 1080
QY 1357 AGCCGGGCGGGGAGAGCTTCACTTTCCTGAGATGATTCATTCGCGCACGGGCT 1416
Db 1081 AGCCGGGCGGGGAGAGCTTCACTTTCCTGAGATGATTCATTCGCGCACGGGCT 1140
QY 1417 ACCAGAACCGGATTTACACAACAGTGAATAATAGATCAACAACCTTTTGGAAAGATTTCC 1476
Db 1141 ACCAGAACCGGATTTACACAACAGTGAATAATAGATCAACAACCTTTTGGAAAGATTTCC 1200
QY 1477 GGGATTAATTAAGACAGATCTACACGGGCTGTGACATGAGCCGCTGACCCCTTGCCCA 1536
Db 1201 GGGATTAATTAAGACAGATCTACACGGGCTGTGACATGAGCCGCTGACCCCTTGCCCA 1260

QY 1537 ACGACTGCGCGGCTCGAGATCGTGCCCGGGGGCCCGCTACGCGATGAGCGGCTCTTCC 1596
Db 1261 ACGACTGCGCGGCTCGAGATCGTGCCCGGGGGCCCGCTACGCGATGAGCGGCTCTTCC 1320
QY 1597 TGCAGAACCAATTGTTGAGACAACTACCGCAAGGCCCGCTTCACTCCGGGCGCGGGCGG 1656
Db 1321 TGCAGAACCAATTGTTGAGACAACTACCGCAAGGCCCGCTTCACTCCGGGCGCGGGCGG 1380
QY 1657 GCCCGGGGCGGGATACGGAACCGGAGCGGCGGACACCAACGGGCTGTGTGCGCGAGC 1716
Db 1381 GCCCGGGGCGGGATACGGAACCGGAGCGGCGGACACCAACGGGCTGTGTGCGCGAGC 1440
QY 1717 AGCCGAGAGACCCGGGCGCGCTCGCGGCAACCTGATTTGTGACGCGCGGCAACACC 1776
Db 1441 AGCCGAGAGACCCGGGCGCGCTCGCGGCAACCTGATTTGTGACGCGCGGCAACACC 1500
QY 1777 GGGTGAACCTTGGGAGCTCGGCTATGACAGGCGACGGAATTTGGCGGCAACCGGCGCA 1836
Db 1501 GGGTGAACCTTGGGAGCTCGGCTATGACAGGCGACGGAATTTGGCGGCAACCGGCGCA 1560
QY 1837 GGTGTCTCTTACCGGGCGGGCGGTGAAGGGGCTGCGGCTGACAGGCTGACAGCTGCA 1896
Db 1561 GGTGTCTCTTACCGGGCGGGCGGTGAAGGGGCTGCGGCTGACAGGCTGACAGCTGCA 1620
QY 1897 CTGGCCCGCGCTCGGCTACTACGCGGACCCGTCGGGCTGGGGCGCGGCAATCCCGC 1956
Db 1621 CTGGCCCGCGCTCGGCTACTACGCGGACCCGTCGGGCTGGGGCGCGGCAATCCCGC 1680
QY 1957 AGTACTGGGAGCAAGTGGGCTGTGGTCTGCTGTGCGCCACAGCGCGCGCGCG 2016
Db 1681 AGTACTGGGAGCAAGTGGGCTGTGGTCTGCTGTGCGCCACAGCGCGCGCGCGCG 1740
QY 2017 CCGGGCGGATGCGGGCGGCGCAATCCCTACCTGGGGGAGAGGCGGAGGCGTGGCGCGG 2076
Db 1741 CCGGGCGGATGCGGGCGGCGCAATCCCTACCTGGGGGAGAGGCGGAGGCGTGGCGCGG 1800
QY 2077 AGCGCTCGCGCTCGCGCGCGCGCGCGGAGGAGCGGCAAGGCACTGTCCGATT 2136
Db 1801 AGCGCTCGCGCTCGCGCGCGCGCGCGCGGAGGAGCGGCAAGGCACTGTCCGATT 1860
QY 2137 CCAAGTGAATGAGACGCGCTCTGTGATCAAGTCAATGACTCAAGGACTCTGGGATTT 2196
Db 1861 CCAAGTGAATGAGACGCGCTCTGTGATCAAGTCAATGACTCAAGGACTCTGGGATTT 1920
QY 2197 ACGAGAGGCGCAAGCGGAGGCGGATCTGCGCGGCGGACAGCGCGGTGCGGAGTTTCT 2256
Db 1921 ACGAGAGGCGCAAGCGGAGGCGGATCTGCGCGGCGGACAGCGCGGTGCGGAGTTTCT 1980
QY 2257 CCGGCTCAAGAGGAGGTGTGCGCGGAGCTGCGGAGCTGCGGAGAACTGCGCCAAAGCA 2316
Db 1981 CCGGCTCAAGAGGAGGTGTGCGCGGAGCTGCGGAGCTGCGGAGAACTGCGCCAAAGCA 2040
QY 2317 TTAAGGGCTACTATGAGCTTCTACTGCAACGTAAGCGCGCTGCGCGCGCGCGCG 2376
Db 2041 TTAAGGGCTACTATGAGCTTCTACTGCAACGCTAGGCGCGCGCTTACCGCGCGCGCG 2100
QY 2377 GCGGGCGGGAGCCCCGAGCGGCGGCTTCCAGAGCTTCCAGGCTGCGGCTGCGGCACT 2436
Db 2101 GCGGGCGGGAGCCCCGAGCGGCGGCTTCCAGAGCTTCCAGGCTGCGGCTGCGGCACT 2160
QY 2437 CCTCCTTGGCAGACCACTCATTTTATTTGACCTCTGATGAGCGGCTGACAGGAATAAGTG 2496
Db 2161 CCTCCTTGGCAGACCACTCATTTTATTTGACCTCTGATGAGCGGCTGACAGGAATAAGTG 2220
QY 2497 CAGGTCTCGAGCGTGAATTTTAACTTTTGTGACACAGAGTCTGTGCAATTAAGTCAACG 2556
Db 2221 CAGGTCTCGAGCGTGAATTTTAACTTTTGTGACACAGAGTCTGTGCAATTAAGTCAACG 2280
QY 2557 ACCTTCAACTTTTGTGTAAACCTTTTGTGTTTCTTACTTACTCTTCTGTGTGAGATTAT 2616
Db 2281 ACCTTCAACTTTTGTGTAAACCTTTTGTGTTTCTTACTTACTCTTCTGTGTGAGATTAT 2340
QY 2617 CCTCTCAAAATTTCCCTCCCTCGTCTTCTTACTCTTACTCTTCTTCTTGTAAAT 2676

QY	939	CTTTGGCTGAAATTTTCA	CCGGGCA	CCAAACGGAGATGATCATCA	CCAAACGGGAGAGCGC	998
Db	661	CTTTGGCTGAAATTTTCA	CCGGGCA	CCAAACGGAGATGATCATCA	CCAAACGGGAGAGCGC	720
QY	999	ATGTTTCCCTTTTAA	GTTTTAA	CATTTCCGGCTCGATCCCA	CGGCTCATTAACAATATT	105
Db	721	ATGTTTCCCTTTTAA	GTTTTAA	CATTTCTGGCTCGATCCCA	CGGCTCATTAACAATATT	780
QY	1059	TTTGTGATGTGATTTT	TGGCGAT	CCCAATCACTGAGGTTTCA	GAGGCAATGGGTT	111
Db	781	TTTGTGATGTGATTTT	TGGCGAT	CCCAATCACTGAGGTTTCA	GAGGCAATGGGTT	840
QY	1119	CCTTGGGCAAAAGCGG	ACCAATG	GCAGAGAAATCGGGCTCATATG	CATCCGAAATTC	117
Db	841	CCTTGGGCAAAAGCGG	ACCAATG	GCAGAGAAATCGGGCTCATATG	CATCCGAAATTC	900
QY	1179	CCCAACACTGGGGCT	CACCTGAT	GCACCAAGAAATCTTTT	GAAAAATTAAACTTACG	123
Db	901	CCCAACACTGGGGCT	CACCTGAT	GCACCAAGAAATCTTTT	GAAAAATTAAACTTACG	960
QY	1239	AACAACAAGAGACTT	CAATATACAT	TGGCAGATGGTGTTTACAGT	CTCTTGCAACG	129
Db	961	AACAACAAGAGACTT	CAATATACAT	TGGCAGATGGTGTTTACAGT	CTCTTGCAACG	1020
QY	1299	TACCAAGCCCCGCT	GATGTGTG	GAAGTAAGTAAGAGACGGGCA	CGGAGACATACGACG	135
Db	1021	TACCAAGCCCCGCT	GATGTGTG	GAAGTAAGTAAGAGACGGGCA	CGGAGACATACGACG	108
QY	1359	CCCGGCGCGGTGAG	ACGTTCACTTT	CCTGAGACTCATGTTCAAT	TCGCGTCAACCGCTTAC	141
Db	1081	CCCGGCGCGGTGAG	ACGTTCACTTT	CCTGAGACTCATGTTCAAT	TCGCGTCAACCGCTTAC	114
QY	1419	CAGAACACGGATTT	ACACAACG	AAATAGATCACAA	CCCTTTTGCAAAAGATTTCCG	147
Db	1141	CAGAACACGGATTT	ACACAACG	AAATAGATCACAA	CCCTTTTGCAAAAGATTTCCG	120
QY	1479	GATTAATTATGACA	CGATCTACAC	CGGCTGTGACATGGA	CCGCTGACCCCTCGCCCAAC	153
Db	1201	GATTAATTATGACA	CGATCTACAC	CGGCTGTGACATGGA	CCGCTGACCCCTCGCCCAAC	126
QY	1539	GACTCGCCCGCTCG	GAGATCGTG	CCCCGGGGCCCCGCTACACG	CATGGCCGGGCTCTTTCTCG	159
Db	1261	GACTCGCCCGCTCG	GAGATCGTG	CCCCGGGGCCCCGCTACACG	CATGGCCGGGCTCTTTCTCG	132
QY	1599	CAGGACCAATTCGT	GAGCAACTAC	CGGCAAGGCCCCGTTTCA	CCCGGCGCGGGCGCGGGC	165
Db	1321	CAGGACCAATTCGT	GAGCAACTAC	CGGCAAGGCCCCGTTTCA	CCCGGCGCGGGCGCGGGC	138
QY	1659	CCCCGGGCGGGGTA	CGGACCGG	CAGCGTGCGGCA	CAACCAACGGGCTGCTGTGCGCGACG	171
Db	1381	CCCCGGGCGGGGTA	CGGACCGG	CAGCGTGCGGCA	CAACCAACGGGCTGCTGTGCGCGACG	144
QY	1719	GCGGAGAGCCGGG	GCAGCCCT	CGCGCGCAACGCTGGTTT	TGACGCGCGCAACACCGG	177
Db	1441	GCGGAGAGCCGGG	GCAGCCCT	CGCGCGCAACGCTGGTTT	TGACGCGCGCAACACCGG	150
QY	1779	CTGGAAGTTCGGG	CTTCGACTAT	GCAGCGGCA	CGGACTTTCGCGGCAACGCGGCCACG	183
Db	1501	CTGGAAGTTCGGG	CTTCGACTAT	GCAGCGGCA	CGGACTTTCGCGGCAACGCGGCCACG	156
QY	1839	CTGGCTCTTTTAC	CGCGCGCGGCGG	CGTAAGGCGCTCGCTG	AGAGCTGAGAGCTGCACT	189
Db	1561	CTGGCTCTTTTAC	CGCGCGCGGCGG	CGTAAGGCGCTCGCTG	AGAGCTGAGAGCTGCACT	162
QY	1899	GCGCGGCGCGCT	CGGCTACTAC	AGCGCGACCCCTG	CGGCTGAGGCGCGCGCACTGCCCGCAG	199
Db	1621	GCGCGGCGCGCT	CGGCTACTAC	AGCGCGACCCCTG	CGGCTGAGGCGCGCGCACTGCCCGCAG	168
QY	1959	TACTGCGGACCA	AGTCGGGCTGGT	GTGTGCTGCTG	CGCCCAACAGCGCGCGGCGCC	201
Db	1681	TACTGCGGACCA	AGTCGGGCTGGT	GTGTGCTGCTG	CGCCCAACAGCGCGCGGCGCC	174

QY	2019	GC	GC	G	C	A	T	G	C	C	G	C	C	A	A	T	C	C	T	A	C	T	G	G	G	C	C	A	G	A	G	G	C	C	A	G	G	C	C	T	G	C	C	G	A		2078		
Db	1741	GC	GC	G	C	A	T	G	C	C	G	C	C	A	A	T	C	C	T	A	C	T	G	G	G	C	C	A	G	A	G	G	C	C	A	G	G	C	C	T	G	C	C	G	A		1800		
QY	2079	CG	CT	G	C	G	C	T	G	C	C	G	G	C	C	G	G	C	C	A	G	A	G	C	C	A	A	G	C	A	A	G	C	A	G	A	C	T	G	C	C	A	T	T	C		2138		
Db	1801	CG	CT	G	C	G	C	T	G	C	C	G	G	C	C	G	G	C	C	A	G	A	G	C	C	A	A	G	C	A	A	G	C	A	G	A	C	T	G	C	C	A	T	T	C		1860		
QY	2139	AG	CT	G	A	T	C	G	A	G	C	G	C	C	T	C	T	C	G	A	T	C	A	G	T	C	A	G	C	A	G	C	A	G	C	A	G	C	A	T	G	C	G	A	T	T	A		2198
Db	1861	AG	CT	G	A	T	C	G	A	G	C	G	C	C	T	C	T	C	G	A	T	C	A	G	T	C	A	G	C	A	G	C	A	G	C	A	G	C	A	T	G	C	G	A	T	T	A		1920
QY	2199	G	A	G	C	A	G	C	C	A	G	C	C	G	A	G	C	C	G	A	T	T	C	G	C	C	G	C	A	C	A	C	G	C	C	C	G	T	C	C	G	A	G	T	T	C		2258	
Db	1921	G	A	G	C	A	G	C	C	A	G	C	C	G	A	G	C	C	G	A	T	T	C	G	C	C	G	C	A	C	A	C	G	C	C	C	G	T	C	C	G	A	G	T	T	C		1980	
QY	2259	CC	GC	T	C	A	A	G	A	C	G	A	G	G	T	G	C	C	C	A	G	C	C	C	C	A	G	C	C	A	G	A	A	C	T	G	C	C	C	A	A	G	A	C	A	T		2318	
Db	1961	CC	GC	T	C	A	A	G	A	C	G	A	G	G	T	G	C	C	C	A	G	C	C	C	C	A	G	C	C	A	G	A	A	C	T	G	C	C	C	A	A	G	A	C	A	T		2040	
QY	2319	AG	CG	G	C	T	A	T	G	C	C	T	C	T	A	C	T	C	G	A	G	A	G	C	G	C	C	C	C	T	G	C	C	C	C	C	C	C	C	C	C	C	C	C		2378			
Db	2041	AG	CG	G	C	T	A	T	G	C	C	T	C	T	A	C	T	C	G	A	G	A	G	C	G	C	C	C	T	A	C	G	C	C	C	C	C	C	C	C	C	C	C	C		2100			
QY	2379	CG	G	C	C	C	C	G	A	C	C	C	C	C	C	T	C	A	C	A	G	C	T	T	T	C	C	C	C	A	G	C	T	T	C	C	C	C	C	C	C	C	C	C		2438			
Db	2101	CG	G	C	C	C	C	G	A	C	C	C	C	C	C	C	T	C	A	C	A	G	C	T	T	T	C	C	C	C	A	G	C	T	T	C	C	C	C	C	C	C	C	C		2160			
QY	2439	T	C	C	T	T	G	C	C	A	C	C	C	T	A	T	T	T	A	T	T	T	A	C	C	T	G	A	T	G	C	C	T	G	C	A	G	A	T	A	G	C	A		2498				
Db	2161	T	C	C	T	T	G	C	C	A	C	C	C	T	A	T	T	T	A	T	T	T	A	C	C	T	G	A	T	G	C	C	T	G	C	A	G	A	T	A	G	C	A		2220				
QY	2499	G	G	T	C	C	G	A	G	C	G	A	T	T	T	A	C	T	T	T	T	T	G	C	A	C	A	G	A	G	C	A	G	A	T	T	A	C	T	G	C	A		2558					
Db	2221	G	G	T	C	C	G	A	G	C	G	A	T	T	T	A	C	T	T	T	T	T	G	C	A	C	A	G	A	G	C	A	G	A	T	T	A	C	T	G	C	A		2280					
QY	2559	C	T	T	C	A	C	T	T	T	G	A	A	C	T	T	T	T	T	T	C	T	A	C	T	T	C	T	T	C	T	G	T	A	T	T	C		2618										
Db	2281	C	T	T	C	A	C	T	T	T	G	A	A	C	T	T	T	T	T	T	T	C	T	A	C	T	T	C	T	T	C	T	G	T	A	T	T	C		2340									
QY	2619	T	C	C	T	A	C	A	T	T	C	C	C	C	C	C	T	C	T	T	T	C	T	T	A	C	C	T	C	T	C	T	T	T	T	T	T	T	T	T	T	T	T		2678				
Db	2341	T	C	C	T	A	C	A	T	T	C	C	C	C	C	C	C	T	C	T	T	T	C	T	T	A	C	C	T	C	T	T	T	T	T	T	T	T	T	T	T	T		2398					
QY	2679	A	A	C	T	T	C	A	C	T	T	A	G	A	G	C	T	T	G	A	G	A	C	T	T	C	A	G	A	G	C	A	G	A	T	T	C	C	A	C	C	C	C		2738				
Db	2399	A	A	C	T	T	C	A	C	T	T	A	G	A	G	C	T	T	G	A	G	A	C	T	T	C	A	G	A	G	C	A	G	A	T	T	C	C	A	C	C	C	C		2455				
QY	2739	A	A	G	T	C	T	G	G	-	C	T	C	A	C	A	T	T	A	C	A	T	A	G	A	T	T	G	A	C	T	T	A	C	T	T	A	C	T	T	A	C		2797					
Db	2456	A	A	G	T	C	T	G	G	-	C	T	C	A	C	A	T	T	A	C	A	T	A	G	A	T	T	G	A	C	T	T	A	C	T	T	A	C	T	T	A	C		2515					
QY	2798	G	G	C	T	C	T	T	T	A	T	C	C	C	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A		2857				
Db	2516	G	G	C	T	C	T	T	T	A	T	C	C	C	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A		2574				
QY	2858	T	T	T	A	G	T	A	G	A	C	C	T	G	A	A	T	G	A	T	T	G	A	C	T	A	G	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T		2917				
Db	2575	T	T	-	-	-	A	G	T	A	G	A	C	C	T	G	A	A	T	G	A	T	T	G	A	C	T	A	G	T	T	T	T	T	T	T	T	T	T	T	T	T		2628					
QY	2918	C	T	A	G	T	T	T	T	A	T	T	C	G	T	C	A	A	G	C	A	A	A	C	A	A	A	C	A	A	A	C	A	A	A	C	A	A	A	C	A	A	A		2977				
Db	2629	C	-	A	G	T	T	T	T	A	T	T	C	G	T	C	A	A	G	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		2685					
QY	2978	T	T	T	C	T	T	T	G	C	T	T	C	T	C	T	C	A	A	T	T	T	C	T	C	T	C	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T		3037					
Db	2686	T	T	T	C	T	T	T	G	C	T	T	C	T	C	C	T	C	A	A	T	T	T	C	T	C	T	C	T	C	T	T	T	T	T	T	T	T	T	T	T	T		2745					
QY	3038	T	G	A	G	T	A	T	A	T	T	C	T	A	G	G	C	T	T	A	A	A	C	T	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		3097					
Db	2746	T	G	A	G	T	A	T	A	T	T	C	T	A	G	G	C	T	T	A	A	A	C	T	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		2805					
QY	3098	C	T	T	C	C	C	A	T	C	G	T	T	C	A	T	A	T	T	A	T	T	A	C	T	A	T	T	A	T	T	A	T	T	A	T	T	A	T	T	A	T	T		3157				

DB 2806 CTTCCCTCAATCCCTTGATGATAAATAATTAATGATGCTTAATGACACAAATAGCTA 2865
 3158 AGGAGATTCACCCCAACACCTTTAAAGG 3186
 DB 2866 AGGAGATTCACCCCAACACCTTTAAAGG 2894

RESULT 3
 BC052737 3992 bp mRNA linear ROD 06-JUL-2005
 LOCUS Mus musculus T-box brain gene 1, mRNA (CDNA clone MGC:64688
 DEFINITION IMAGE:6825394), complete cds.

ACCESSION BC052737
 VERSION BC052737.1 GI:30931365
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 1 (bases 1 to 3992)
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diachenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 Stalcron M., Soares M.B., Bonaldo M.F., Casavant T.L.,
 Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
 Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
 Abramson R.D., Mullany S.J., Bosak S.A., McKwan P.J.,
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
 Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettman M., Madan A., Rodriguez S.,
 Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,
 Bouffard G., Blakeley R.W., Touchman J.W., Green E.D.,
 Dickson M.C., Rodriguez A.C., Grimwood J., Schultz J., Myers R.M.,
 Butterfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E.,
 Scherck A., Schein J.E., Jones S.J. and Marra M.A.
 Mammalian Gene Collection Program Team
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 3992)

CONSRMT
 TITLE NIH MGC Project
 JOURNAL Direct Submission
 PUMED Submitted (19-MAY-2003) National Institutes of Health, Mammalian
 REFERENCE Gene Collection (MGC), Bethesda, MD 20892-2590, USA
 AUTHORS NIH-MGC Project URL: http://mgc.nhl.nih.gov
 COMMENT Contact: MGC help desk
 Email: cgabds-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant.
 Web site: http://genome.uiowa.edu
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo M.F., Akabogu I., Bair T., Bair J., Crouch K., Davis A.,
 Fisher K., Keppel C., Kucaba T., Lebeck M., Melo A., Schaefer K.,
 Schetz T., Smith C., Snir E., Tack D., Trout K., Walters J.,
 Casavant T., Soares M.B.

FEATURES
 SOURCE
 1. 3992

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILML at: http://image.llnl.gov
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 34328150.
 Location/Qualifiers

gene
 1. 3992
 /gene="Tbr1"
 /db_xref="GeneID:21375"
 /db_xref="MGI:107404"
 256_2301
 /gene="Tbr1"
 /codon_start=1
 /product="T-box brain gene 1"
 /protein_id="AAH52737.1"
 /db_xref="GI:30931365"
 /db_xref="GeneID:21375"
 /db_xref="MGI:107404"
 /translation="MQLHECLSPSILMSKRFVNSSYPHSGSEIVLHDPITSTD
 NLERSSPLKTRGMTNQSDTDFPDSKDSFVQSKLSFVLGVSELRHSFPGSAA
 DRYLLSQSSOPGSAATAPSAFPPYSGHPAPHAFISGSPRYMAHHPVITNGATNSL
 LNSSPGQYFAGYPPQOYGHYOGAPFYQFSTQGLVPGKQVYLCNPLDKFL
 RHOMETITKQGRMPPPLSPNISGLDPTAYNLFVYIADPNHMFQGGKVPCKG
 ADTVQGNRYVNHDSPTNGAHMMROEISRKLTLNNKASNNNGQVYVQSLHKYQ
 PRLHVEVNEDETDTSOPGVQFTTPPETPFIATVATVQNTDIQKIDHNPAPKGR
 DNYDTIITGCMNDLTPSPNDSPSSQIVPGARVMAQSPLODQVSNYAKRFRPGAG
 AGPGRDPSVPHNGLSPQADPDGAPVSPORFVPPANNRLDFPASAQSVLPFAG
 NATULLSAGVAVALPLQAACTGRPLGYADPSGWCARSPQYCGAKSGSVLPCW
 NSAAABRMAGAVLGEAEGLAABRSPLAPALDEKPKLSSSTETSSISKID
 SSDGIYEQARRRISPDITVSSSSSLKSEVLAQNDCEKNACKIGYVGFPSHS"

ORIGIN
 Query Match 78.2%; Score 2490; DB 9; Length 3992;
 Beet Local Similarity 90.0%; Pred. No. 2,8e-311;
 Matches 2870; Conservative 0; Mismatches 245; Indels 73; Gaps 16;

DB 1 CAGCTAACCAATGGAGAGCGCTTTCATCGCGTGCACAGACGAGAAATCAAAAGTAG 60
 QY 47 CAGCTAACCAATGGAGAGCGCTTTCATCGCGTGCACAGACGAGAAATCAAAAGTAG 106
 DB 1 CAGCTAACCAATGGAGAGCGCTTTCATCGCGTGCACAGACGAGAAATCAAAAGTAG 60
 QY 107 CAGCTAACCAATGGAGAGCGCTTTCATCGCGTGCACAGACGAGAAATCAAAAGTAG 166
 DB 61 CAGCTAACCAATGGAGAGCGCTTTCATCGCGTGCACAGACGAGAAATCAAAAGTAG 120
 QY 167 GAAATTCACCACTTGAAGAGATTTGCTGTTGAAGTCTTCTGTAGTGAAGGAGT 226
 DB 121 GAAATTCACCACTTGAAGAGATTTGCTGTTGAAGTCTTCTGTAGTGAAGGAGT 179
 QY 227 CTTGATGATTTCTAGTTATGATTAATAGACCTTTAAAAACGAGGACGGAGCGGAGT 286
 DB 180 CTTGATGATTTCTAGTTATGATTAATAGACCTTTAAAAACGAGGACGGAGCGGAGT 239
 QY 287 TTTGAGTTTAAAGCTATGACAGTGGAGACCTGCTTCTCTCTTATCATGCTCCAA 346
 DB 240 TTTGAGTTTAAAGCTATGACAGTGGAGACCTGCTTCTCTCTTATCATGCTCCAA 299
 QY 347 GAAATTCATGATGAGAGAGAGCTACCACTTCAAGCGAGATCCGACTTGTCTTGA 406
 DB 300 GAAATTCATGATGAGAGAGAGCTACCACTTCAAGCGAGATCCGACTTGTCTTGA 359
 QY 407 CGATCATCCCATATCTCGACCACTGACCACTGAGAGAGACTTCACTTTAAAAAAT 466
 DB 360 TGATCATCCCATATCTCGACCACTGACCACTGAGAGAGACTTCACTTTAAAAAAT 419
 QY 467 TACAGGGGAG 526
 DB 420 TACAGGGGAG 479
 QY 527 AGGGAGCGTCCAG 586

Db 480 AGGGAGCGTCCAGAGAAGTAAACTCTCTCTGTCTGAGCGGGGTCTCTGAGCTTCGTCA 539
Qy 587 CAGTTTCATGAGCTCTGCTGAGATGCTACTCTCTCTCTCATGTCACAGCCAGCTC 646
Db 540 CAGTTTCATGAGCTCTGCTGAGATGCTACTCTCTCTCATGTCACAGCCAGCTC 599
Qy 647 TGGGGCCACTGCTCCAGTGGCCATGTTCCGTCACCCGGCCAGCAAGCAAGCGGCAACC 706
Db 600 TGGGGCCACTGCTCCAGTGGCCATGTTCCGTCACCCGGCCAGCAAGCAAGCGGCAACC 659
Qy 707 CGCCTTCTCATCGGACAGCCCTTACGCGCTACATGAGCCCAACCCCGGTCAATCAACCAAG 766
Db 660 CGCCTTCTCATCGGACAGCCCTTACGCGCTACATGAGCCCAACCCCGGTCAATCAACCAAG 719
Qy 767 AGCTTACAACAGCT 826
Db 720 AGCTTACAACAGCT 779
Qy 827 CTACCCACAGAGTACGCGCACTCTCTACCAAGAGCTCGTTCTACAGATTCTCTCTCAC 886
Db 780 CTACCCACAGAGTACGCGCACTCTCTACCAAGAGCTCTCTCTCTCTCTCTCTCTCTCAC 839
Qy 887 CCAGCCGAGGCTGTGTGCGCGGCAAGCAAGGTGTACTCTGTGCAACAGGCCCTTTGGCT 946
Db 840 CCAGCCGAGGCTGTGTGCGCGGCAAGCAAGGTGTACTCTGTGCAACAGGCCCTTTGGCT 899
Qy 947 GAAATTTTCAACCGGCAACCAACGGAATGATCATCAACCAACAGGGAAGCGCATTTTCC 1006
Db 900 GAAATTTTCAACCGGCAACCAACGGAATGATCATCAACCAACAGGGAAGCGCATTTTCC 959
Qy 1007 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1066
Db 960 CTTTGTGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1019
Qy 1067 TGTGATTTTGGCGGATCCCAATCACTGTGAGGTTTCAAGAGGCAATGAGTTCTTGGCG 1126
Db 1020 TGTGATTTTGGCGGATCCCAATCACTGTGAGGTTTCAAGAGGCAATGAGTTCTTGGCG 1079
Qy 1127 CAAAGCGGACCAATGTGCAAGAAATGGGTCTATATGCAATCCGGAATTTCCCAACAC 1186
Db 1080 CAAAGCGGACCAATGTGCAAGAAATGGGTCTATATGCAATCCGGAATTTCCCAACAC 1139
Qy 1187 TGGGGCTCATGAGTGGCGCAAGAAATCTCTTTGAAATTTAAATTTAAATTTAAATTTAA 1246
Db 1140 TGGGGCTCATGAGTGGCGCAAGAAATCTCTTTGAAATTTAAATTTAAATTTAAATTTAA 1199
Qy 1247 AGGAGCTTCAAAATTAACAATGAGGAGATGTTGTTTAAAGTCTTGTCAAAATACAGACC 1306
Db 1200 AGGAGCTTCAAAATTAACAATGAGGAGATGTTGTTTAAAGTCTTGTCAAAATACAGACC 1259
Qy 1307 CCGCTCGCATGTGTGGAAGTGAACGAGACGGCAACGAGACCTAGCCAGCCCGCG 1366
Db 1260 CCGCTCGCATGTGTGGAAGTGAACGAGACGGCAACGAGACCTAGCCAGCCCGCG 1319
Qy 1367 CGTGGAGAGCTCATCTTTCCCTGAGACCTGCTTCCGCGTCAACGCGCTACCAAGAC 1426
Db 1320 AGTCCAGAGTTTCACTTTCCGAGACTAGTTTCACTGCTGTCAACGCGCTACCAAGAC 1379
Qy 1427 GGATATTAACAACCTGAAATATGATCAACACCTTTTGTCAAAAGATTTCCGGATTAATTA 1486
Db 1380 GGATATTAACAACCTGAAATATGATCAATACCTTTTGTCAAAAGATTTCCGATTAATTA 1439
Qy 1487 TGAACAGATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCAACGATCGCC 1546
Db 1440 TGAACAGATCTACACCGGCTGTGACATGAGACCGCTGACCCCTGCGCAACGATCGCC 1499
Qy 1547 GGGCTCGAGATCGTGGCCCGGGGCGCGTACAGCANTGGCGGCTCTTTCTGTGAGACCA 1606
Db 1500 GGGCTCGAGATCGTGGCCCGGGGCGCGTACAGCANTGGCGGCTCTTTCTGTGAGACCA 1559
Qy 1607 GTTCTGAGCAACTACGCGCAAGGCGCTTCCACCGGCGCGGCGCGCGCGCGCGCGCG 1666
Db 1560 GTTCTGAGCAACTACGCGCAAGGCGCGCTTCCACCGGCGCGCGCGCGCGCGCGCGCGCG 1619

Qy 1667 GGGTACGGACCGGAGCGTGCAGCAACAAAGGGCTGTGTGCGCCGACAGGCGGAGGA 1726
Db 1620 GGGTACGGACCGGAGCGTGCAGCAACAAAGGGCTGTGTGCGCCGACAGGCGGAGGA 1679
Qy 1727 CCGGGGCGCGCCCTCGCGCGCAACGCTGTGTTTGTGAACGCGCGGCGCAACACCGGCTGA 1786
Db 1680 CCGGGGCGCGCCCTCGCGCGCAACGCTGTGTTTGTGAACGCGCGGCGCAACACCGGCTGA 1739
Qy 1787 CCGGGGCTCGGCTCTATGACACGCGCAACGACCTTCCGCGGCGCAACGCGGCGCAACG 1846
Db 1740 CCGGGGCTCGGCTCTATGACACGCGCAACGACCTTCCGCGGCGCAACGCGGCGCAACG 1799
Qy 1847 TTAACGGGCGGCGGCGTGAAGGCGTGCAGCTGACAGGCTGACAGGCTGACAGGCGGCG 1906
Db 1800 GTATGCGGCGCGGCGGCGTGAAGGCGTGCAGCTGACAGGCGGCGGCGGCGGCGGCGGCG 1859
Qy 1907 GCTTGGCTACTACGCGCAACCGTTCGAGCTGAGGCGCGCGCAAGTCCCGCGAGTACTGCG 1966
Db 1860 GCTTGGCTACTACGCGCAACCGTTCGAGCTGAGGCGCGCGCAAGTCCCGCGAGTACTGCG 1919
Qy 1967 CACCAAGTGGGCTCGGCTGTGCGCTGTGCGCCCAACAGCGCGCGCGCGCGCGCAT 2026
Db 1920 CCGCAAGTGGGCTCGGCTGTGCGCTGTGCGCCCAACAGCGCGCGCGCGCGCGCAT 1979
Qy 2027 GGGCGGCGCAATCCCTACCTGGGCGAGAGGCGCGAGGCGCTGAGCGCGCGCGCGCGCG 2086
Db 1980 GGGCGGCGCAATCCCTACCTGGGCGAGAGGCGCGAGGCGCTGAGCGCGCGCGCGCGCG 2039
Qy 2087 GCTGCGCGCGCGCGCGCGCGAGGAGCGCAAGCGCGCAAGGCTGTCCGATTTCAAGCTGAT 2146
Db 2040 GCTGCG 2096
Qy 2147 CGAAGCGCGCTCTCTCATCAAGTTCATGACTGACAGGACTCGGGGATTTTACAGAGCGC 2206
Db 2097 CGAAGCGCGCTCTCTCATCAAGTTCATGACTGACAGGACTCGGGGATTTTACAGAGCGC 2156
Qy 2207 CAAAGGAGGCGAGTCTCGCGCGCGAGCGCGCGCGGTGCGAGAGTTCGTCGCGGCTCAA 2266
Db 2157 CAAAGGAGGCGAGTCTCGCGCGCGAGCGCGCGCGGTGCGAGAGTTCGTCGCGGCTCAA 2216
Qy 2267 GAGCGAGTGTGCGCGCGCGAGGAGCTGCGAAGAACTGCGCGCAAGAGACTTAAAGCGCTA 2326
Db 2217 GAGCGAGTGTGCGCGCGCGAGGAGCTGCGAAGAACTGCGCGCAAGAGACTTAAAGCGCTA 2276
Qy 2227 CTATGCTTCTTACTGCAAGCTTAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2386
Db 2277 CTATGCTTCTTACTGCAAGCTTAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2336
Qy 2387 ACCCGAGCGAGCGCGCTCAAGCTTTCCGAGCTCCGCGCTCCGCAACCTCTCTTGGCG 2446
Db 2337 GGGCG 2384
Qy 2447 CACCGACTCATTTTATTTTGAACCTCGATGAGCGGCTGTGCAAGCAATAGTCAAGTCTCG 2506
Db 2385 CACCGACTCATTTTATTTTGAACCTCGATGAGCGGCTGTGCAAGCAATAGTCAAGTCTCG 2442
Qy 2507 AGCGTATTTTAACTTTT-----TGACAGCACTCTCTG 2542
Db 2443 ACTATGATTTTAAACCTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTG 2502
Qy 2543 CAATTTAGTCAACGAGCTTCAACCTTTGCTGTAACCTTTTGTGTTTCTTACTTACTTCTG 2602
Db 2503 CAGTTAGGCAACGAGCTTGAACCTGCTGTGAACCTTTGTGTTTCTTCAACTT-----TTC 2558
Qy 2603 TTTCTGAGATTAATCTCTTAC-----AATTCCTCTCCCTCTGTCTTTCTTTACTCTC 2656
Db 2559 GTCTGTAGATTAAGTAACTCTCTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2618
Qy 2657 CTACTTCTCTTTCTTTGTAATGAACCTTTTCACTTTTGAAGACTGTGGGCACTCTGTACG 2716
Db 2619 CCTCTCTTTCTTTGTAATGAACCTTTTCACTTTTGAAGACTGTGGGCAATCTGTACG 2678

Db	1	CTCTTCTCCACAGGAAAAAAGTAATCAAACCTTTGAGAGCAATTTGCTGGTTGAA	60
Qy	203	GTGCTTTCTGTCTAGTGAAGGGGCTGTGAGATTTTCTAGTTATGATTAATAAGACCTTTAA	262
Db	61	GTGCTTTCTGTCTTAGTGA - GGGGCTGTGAGATTCTTAGTGTATGATTAACAGACTTTAA	119
Qy	263	AAACCAAGGAGCGGGAGGGCGAGGTTCAGGTTCTAGAGCTATGACGCTGAGCACTGGCT	322
Db	120	AAACCAAGGAGCGGGAGGGCGACGTTCAAGTTCAGAGCTAATGACGCTGAGCACTTGGCT	179
Qy	323	TTCTCCTTTATCAGCTCTCCAAAGAAATTTCTCAATGAGACACAGCTACCCACATTC	382
Db	180	CTCTCCTTTATCAGCTCTCCAAAGAAATTTCTCAATGAGACACAGCTACCCACATTC	239
Qy	383	AGCGGATCCGAGCTTGTCTTGCACGATCATCCATTAATCTGCACACTGACAACTGGA	442
Db	240	GGGGCGGATCTGAGCTTGTCTTGCATGATCATCCATTAATCTGCACACTGACAACTGGA	299
Qy	443	GAGAACTTCACTTTGAAAAAATTAACAAGGGGATGACGAATCAGTCAATACAGACAA	502
Db	300	GAGAACTTCACTTTGAAAAAATTAACAAGGGGATGACGAATCAGTCAATACAGACAA	359
Qy	503	TTTTCTGCTCCAAAGGACTCAACAGGGGAGCTCCAGAGAAATTAATCTCTCCGCTGT	562
Db	360	TTTTCTGCTCCAAAGGACTCAACAGGGGAGCTCCAGAGAAATTAATCTCTCTCTGCTT	419
Qy	563	GGACGGGGTCTGTGAGCTTCGTCAAGTTTCGATGAGCTCTGCTGAGATTCGCTACTCT	622
Db	420	GGACGGGGTCTGTGAGCTTCGTCAAGTTTCGATGAGCTCTGCTGAGATTCGCTACTCT	479
Qy	623	CTCTCAATCAGCCAGCCAGCTCTGCGGCACTGTCTCCAGTGCATGTTCCGCTACC	682
Db	480	CTCTCAATCAGCCAGCCAGCTCTGCGGCACTGTCTCCAGTGCATGTTCCGCTACC	539
Qy	683	GGGCGAGCAGGAGCGGGGCAACCCGGCTTCATTCGGGAGCCCTAGCGGCTCATGAGC	742
Db	540	CAGCCAGCAGGAGCGGGCGATCCGGCTTCTCAATCGGCAACCCAGTGCATATAGGC	599
Qy	743	CCACGACCCGGTATCACCAAGGAGCTTACCAAGAGCTCTGTCCAACTCTCTCGCGCA	802
Db	600	CCACGACCCGGTATTAACCAAGGAGCTTACCAAGAGCTCTGTCCAACTCTCTCGCGCA	659
Qy	803	GGGATACCCACGCGCGGCTACCCCTACACAGATGACGGCACTCTTACAAAGAGC	862
Db	660	GGGATACCCACGCGCGGCTACCCCTACACAGATGACGGCACTCTTACAAAGAGC	719
Qy	863	TCCGTTTACAGATTCTCTCAACCAAGCGGGGCTGGTCCCGGCAAGCAAGATGTA	922
Db	720	CCCTTCTTACAGATTCTCTCAACCAAGCGGGGTTGGTCCCGGCAAGCGCAAGTATA	779
Qy	923	CCTGTGCAACAGAGCCCTTGGCTGAATTTACCGGCAACAAAGAGATGATCATAC	982
Db	780	CCTGTGCAACAGAGCCCTTGGCTGAATTTATGCGGATCAAAAGAGATGATCATAC	839
Qy	983	CAACAGGAGAGCGCATTTTCTTTTAAATTTAAATTTCTGCTTCGATCCAC	1042
Db	840	TAAACAGGAGAGCGCATTTTCTTTTAAATTTAAATTTCTGCTTCGATCCAC	899
Qy	1043	GGCTCATTAATTTTTGTGAGATGTGATTTTTGGCGATCCCATTCAGAGGTTTCA	1102
Db	900	GGCTCATTAATTTTTGTGAGATGTGATTTTTGGCGATCCCATTCAGAGGTTTCA	959
Qy	1103	AGAGGCAATAGGGTTCCTTGGCGCAACCGGACCAATGTGCAAGAAATCGGATCTTA	1162
Db	960	AGAGGCAATAGGGTTCCTTGGCGCAACCGGACCAATGTGCAAGAAACCGGATCTTA	1015
Qy	1163	TATGATCCGATATCCCAACACTGGGGCTCACTGATGCGCAAGAAATCTTTTGG	1222
Db	1020	TATGATCCGATATCCCAACACTGGGGCTCACTGATGCGCAAGAAATCTTTTGG	1072
Qy	1223	AAAAATTAATCTTAGACAAACAAAGAGCTTCAATATAATGCGCAGATGTGTGTTTT	1282
Db	1080	AAAAATTAATCTTAGACAAACAAAGAGCAATCAACATGTGGCAGATGTGTGTTTT	1139

[illegible]

Db 300 CCTGATCCAAAGACATCAACAGGAGACGTCAGAGAAAGTAACTCTCTCTGTTCGAC 359
Qy 567 GGGGCTCTGAGACTTCGTACAGTTTCGATGGCTCTGCTGCAGATCGCTACTCTCTCT 626
Db 360 GGGGCTCTGAGACTTCGTACAGTTTCGATGGCTCTGCTGCAGATCGCTACTCTCTCT 419
Qy 627 CAGTCAGGACGACAGATCTGGGCACTGCTCCAGAGGACATTTCCCGGACCCGAC 686
Db 420 CAGTCAGGACGACAGATCTGGGCACTGCTCCAGAGGACATTTCCCGGACCCGAC 479
Qy 687 CAGCACGGACCGGCGACCCCGCTTCTTCATGGCAGGCTTACGCGCTAACGAGCCAC 746
Db 480 CAGCACGGACCGGCGATCCCGCTTCTTCATGGCAGGCTTACGCGCTAACGAGCCAC 539
Qy 747 CACCGGCTATACCAAGGAGACCTAACAGCTCTGTCCAACTTCTGCGCGAGGGA 806
Db 540 CACCGGCTATACCAAGGAGCTTAACAAGCTGTGTCCAACTTCTGCGCGAGGAC 599
Qy 807 TACCCGACGCGGCTACCCCTACCCACAGAGTACGCGCACTCTTACCAAGAGACTCG 866
Db 600 TACCCGACGCGGCTACCCCTACCCACAGAGTACGCGCACTCTTACCAAGAGACTCG 659
Qy 867 TTTCTACAGATTCTCTCCACCCAGCCGCGGCTGTGCTCCGCGAAGCAAGGTGTACTG 926
Db 660 TTTCTACAGATTCTCTCTCCACCCAGCCGCGGCTGTGCTCCGCGAAGGTATATCTG 719
Qy 927 TCGCAACAGGCCCCCTTTGGTGAAATTTACCGGACCAAAAGGAAATGATCATCAAA 986
Db 720 TCGCAACAGGCACTTTGGTGAAATTTACCGGCACTAAACGGAATGATCATCTAAA 779
Qy 987 CAGGAGAGCGGATGTTTCTTTTAAAGTTTAAATTTCTGTCTGTATCCAGGCT 1046
Db 780 CAGGAGAGCGGATGTTTCTTTTAAAGTTTAAATTTCTGTCTGTATCCAGGCT 839
Qy 1047 CATTAACAATTTTTGTGATGTATTTTGGCGGATCCCAATCACTGAGGTTTCAAGA 1106
Db 840 CATTAACAATTTTTGTGATGTATTTTGGCGGATCCCAATCACTGAGGTTTCAAGA 899
Qy 1107 GCGAAATGGGTTCTTTGGCGGAAAGCGACCAATGTGCAAGGAAATGGGTTATATG 1166
Db 900 GCGAAATGGGTTCTTTGGCGGAAAGCGACCAATGTGCAAGGAAATGGGTTATATG 959
Qy 1167 CATCGGATTTCCCGCAACACTGGGGCTCACTGATGCGCCAAAGAAATCTTTTGAAGA 1226
Db 960 CATCGGATTTCCCGCAACACTGGGGCTCACTGATGCGCCAAAGAAATCTTTTGAAGA 1019
Qy 1227 TTTAAACTTAGAACAACAAGAGCTTCAATTAACAATGGCAGATGGTGTAAACG 1286
Db 1020 TTTAAACTTAGAACAACAAGAGGATCAACAACAATGGCAGATGGTGTAAACG 1079
Qy 1287 TCTTTCGACAAAGTACCAAGCCCGCTGTGATGTGTGTGAAGTAAAGAGATGACAAG 1346
Db 1080 TCTTTCGACAAAGTACCAAGCCCGCTGTGATGTGTGTGAAGTAAAGAGATGACAAG 1139
Qy 1347 GACACTAGGACGCGCGCGCTGACAGTTCATCTTCCCTGAGACTGATTCATCGC 1406
Db 1140 GACACTAGGACGCGCGCGGAGTCCAGAGCTTCACTTTCCGAGACTGATTCATCGT 1199
Qy 1407 GTACCGCTACCGAACAAGATATTTACAACTGAAATAATGATCAACCCCTTTTGA 1466
Db 1200 GTACCGCTACCGAACAAGATATTTACAACTGAAATAATGATCAACCCCTTTTGA 1259
Qy 1467 AAAAGATTTTCGGATTAATTAATGACAGATCTACACCGGCTGTGACATGAGCCG 1526
Db 1260 AAAAGATTTTCGGATTAATTAATGACAGATCTACACCGGCTGTGACATGAGCCG 1319
Qy 1527 CCTCGCCCAAGCACTCGCGCGCTGACAGATGTTGCTCCGCGGCGCGCTACCGCATG 1586
Db 1320 CCGTGCCTCAAGCACTCTCGCGCTGACAGATGTTGCTCCGCGGCGCGCTACCGCATG 1379
Qy 1587 GGGCTTTTCTGTGAGGACAGTTTGTGAGCAACTACGCGCAAGGCGGCTTCAACCGG 1646
Db 1380 GGGCTTTTCTGTGAGGACAGTTTGTGAGCAACTACGCGCAAGGCGGCTTCAACCGG 1439

Qy 1647 GCGGCGCGGAGCCCCCGGAGTACGGAACGACAGCTGCGGACAAACGAGGCTGCTG 1706
Db 1440 GCGGCGCGGAGTCCCGGAGCCGAGGACGAGACGCTGCGGACAAACGAGGCTGCTG 1499
Qy 1707 TCGCGGAGGACCGGAGGACCCCGGCGCGCTCTGCGCGCAAGCTGTGTTTGAAGCCG 1766
Db 1500 TCCCGGAGGAGGCGAGGAGACCCCGGCGCGCTGCGCGGAGGCTGTGTACAGCCG 1559
Qy 1767 GCGAACAACCGGCTGGACTTTCGCGGCTTCCGCTTAATGACAGGCGCAAGCACTTCG 1826
Db 1560 GCGAACAACCGGCTGGACTTTCGCGGCTTCCGCTTAATGACAGGCGCAAGCACTTCG 1619
Qy 1827 AACCGGCGCAAGCTGCTCTTTACGCGGAGCGGAGCTGAGAGGCTGCGCTGAGGCT 1886
Db 1620 AACCGGCGCAAGCTGCTCTTTACGCGGAGCGGAGCTGAGAGGCTGCGCTGAGGCT 1679
Qy 1887 GCAAGCTGCACTGGACGCGCGCTTTCGCTACTACGCGCAACCGCTGTGGGCTGGG 1946
Db 1680 GCGGCTGCAAGGAGCGCGCTGCTGCTACTACTACGCGCAACCGCTGTGGGCTGGG 1739
Qy 1947 AGTCCCCGCAAGTATGCGGCAACCAATCGGAGCTGTGCTGCTGCTGCTGCTGCTG 2006
Db 1740 AGCCCCGCAAGTATGCGGAGCGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1799
Qy 2007 GCGCGGCGCGCGCGCATGCGCGGCAATCCCTACCTGAGGCGAGAGCGGAGGCG 2066
Db 1800 GCGCGGCGCGCGCGCATGCGCGGCAATCCCTACCTGAGGCGAGAGCGGAGGCG 1859
Qy 2067 CTGCGCGCGAGCGCTGCGGCTGCGCGCGCGCGCGCGAGAGACGCGAAGCCCAAG 2126
Db 1860 CTGCGCGCGAGCGCTGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1916
Qy 2127 CTGCGCGATTCAGCTGATGAGACGCGCTCTCTGATGCAAGTCACTGACTCAGGAC 2186
Db 1917 CTGCGCGATTCAGCTGATGAGACGCGCTCTCTGATGCAAGTCACTGACTCAGGAC 1976
Qy 2187 TCGGGAATTTTACGAGCAGGCGCAAGCGGAGCGGATCTGCGCGGCGGACAGCGCG 2246
Db 1977 TCGGGAATTTTACGAGCAGGCGCAAGCGGAGCGGATCTGCGCGGCGGATCTGCT 2036
Qy 2247 GAGAGTTGTCCTCCGCTCAAGAGCGAGGTGTGGCCAGCGGAGCTGCGAAGAACTGC 2306
Db 2037 GAGAGTTGTCCTCCGCTCAAGAGCGAGGTGTGGCCAGCGGAGCTGCGAAGAACTGC 2096
Qy 2307 GCGAAGGACATTTACGCGCTACTATGCTTCTACTCGACAGCTAGCGCGCTGCGCG 2366
Db 2097 GCGAAGGACATTTACGCGCTACTATGCTTCTACTCGACAGCTAGCGCGCTGCGCG 2156
Qy 2367 CCGGCGCGCGCGCGGAGCGCGGAGCGCGGAGCGCGCTCAAGAGCTTTCGCCAGCT 2426
Db 2157 GCGGCGCGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCG 2204
Qy 2427 TCCGCAACATCTCTCTTTCGAGACCCACTCAATTTATTTGACCTTCGATGCGCTG 2486
Db 2205 TTTCCAGGCGCGGCTTTCGAGACCGCTC-TTTCATTTGACCTTCGATGCGCTGCG 2263
Qy 2487 CGAATTAAGTGAAGTCTCCAGCGCTAATTTAACTTTT----- 2526
Db 2264 GG-ATAAGTGAAGTCTCTCACTAATTTTAAACTCTCTTTTCTTTCTTTCTTTCT 2322
Qy 2527 -----TGCAAGAGCTCTCTGCAATTTAGCTCAACGACCTTCAACTTGTGAACCT 2582
Db 2323 TCTCTACAGAGCTTCTCTGAGTTAGGACGACCGACCTTGAACCTGCTGTAAACCT 2382
Qy 2583 GGTTCCTACTTACTCTCTTCTGTGAGATTATCTCTCAATTC-----CCCTGCC 2636
Db 2383 GGTTCCTACTT-----TTGCTGTGAGATTAGATCTCTCTCTTTTCTTTCACACCC 2438
Qy 2637 CTTGCTCTTCTTCTTACTCTCTACTTCTCTTCTTCTTGAATGAACCTTCACTTGA 2696
Db 2439 CTTCTCTTTCGCGCACTCATCTCTCTTCTCTTGTGAATGAACCTTCACTTGA 2498

[illegible]

OY	345	AAAGAAATTTTCATATGAGAGAGAGGATCCCAATTGAGGGGATCCGAGCTGTCTTG	404
Db	121	AAAGAAATTTTCATATGAGAGAGGATCCCAATTGAGGGGATCCGAGCTGTCTTG	180
OY	405	CACGATCATCCCATTTATCTCGACCACTGACGAAAGTTCACTTTGAAAAA	464
Db	181	CACGATCATCCCATTTATCTCGACCACTGACGAAAGTTCACTTTGAAAAA	240
OY	465	ATTACGAGGGGATGACGAAATCAGTCAAGTACAAATTTTCTGACTCCAAAGACTCA	524
Db	241	ATTACGAGGGGATGACGAAATCAGTCAAGTACAAATTTTCTGACTCCAAAGACTCA	300
OY	525	CCAGGGGACGTCCAGAGAAATTAATCTCTCTCTTGGAGGGGGTCTCGAGCTTGT	584
Db	301	CCAGGGGACGTCCAGAGAAATTAATCTCTCTCTTGGAGGGGGTCTCGAGCTTGT	360
OY	585	CACAGTTTCATGAGTCTGTGAGATCGCTACCTCTCTCAGTCCAGCCAGCCACAG	644
Db	361	CACAGTTTCATGAGTCTGTGAGATCGCTACCTCTCTCAGTCCAGCCAGCCACAG	420
OY	645	TCTGGGGCACTGCTCCAGTGCATGTTCCGTAACCCCGGCGAGACGAGCCGGCGAC	704
Db	421	TCTGGGGCACTGCTCCAGTGCATGTTCCGTAACCCCGGCGAGACGAGCCGGCGAC	480
OY	705	CCCGGCTTCTCATATGAGAGCCCTTAAGCCGCTACATGAGCCCAACCGGATATCAACAAC	764
Db	481	CCCGGCTTCTCATATGAGAGCCCTTAAGCCGCTACATGAGCCCAACCGGATATCAACAAC	540
OY	765	GGAGGCTTCAACAAGCTCTGTGCAACTCTCGCGGAGGATACCCACGGCCGGCTTAC	824
Db	541	GGAGGCTTCAACAAGCTCTGTGCAACTCTCGCGGAGGATACCCACGGCCGGCTTAC	600
OY	825	CCCTTACCACAGCAGTACGGGCACTCTTACCAAGAGCTCCGTTTCACTCATGTTCTCTCC	884
Db	601	CCCTTACCACAGCAGTACGGGCACTCTTACCAAGAGCTCCGTTTCACTCATGTTCTCTCC	660
OY	885	ACCCAGCCGGGGCTGTGTGCTCCGGCAACAACAAGATGTAACGTGCAACAGGCCCTTTGG	944
Db	661	ACCCAGCCGGGGCTGTGTGCTCCGGCAACAACAAGATGTAACGTGCAACAGGCCCTTTGG	720
OY	945	CTGAATTTTCAACCGGCAACAAACGAGATGATCATCAACAAACAGGAAAGCGCATGTTT	1004
Db	721	CTGAATTTTCAACCGGCAACAAACGAGATGATCATCAACAAACAGGAAAGCGCATGTTT	780
OY	1005	CCTTTTTTAAGTTTAAATTTCTGATCTCGATCCACGGCTCATTTACAATATTTTGTG	1064
Db	781	CCTTTTTTAAGTTTAAATTTCTGATCTCGATCCACGGCTCATTTACAATATTTTGTG	840
OY	1065	GATGTGATTTTGGCGGATCCCAATCACTGAGAGTTTCAAGAGGCAAAATGGTTCTTGG	1122
Db	841	GATGTGATTTTGGCGGATCCCAATCACTGAGAGTTTCAAGAGGCAAAATGGTTCTTGG	900
OY	1125	GGCAAGCGGAAACCAATGTGCAAGAAATGGGCTTATATCATCCGGAATCCCCCAAC	1182
Db	901	GGCAAGCGGAAACCAATGTGCAAGAAATGGGCTTATATCATCCGGAATCCCCCAAC	960
OY	1185	ACTGGGGGCTCACTGATGCGCCAAAGAAATCTCTTTGGAAAAATTAAACTTACGAACAAC	1244
Db	961	ACTGGGGGCTCACTGATGCGCCAAAGAAATCTCTTTGGAAAAATTAAACTTACGAACAAC	1022
OY	1245	AAAGAGCTTCAAAATTAACAATGGCGAGATGTGTGTTTTCAATCTCTTGGCAAAAGTACGAG	1304
Db	1021	AAAGAGCTTCAAAATTAACAATGGCGAGATGTGTGTTTTCAATCTCTTGGCAAAAGTACGAG	1080
OY	1305	CCCGGCTGATGTGGTGAAGTGAACGAGACGGGACGGAGGACACTAGCCAGCCCGGAC	1366
Db	1081	CCCGGCTGATGTGTGGTGAAGTGAACGAGACGGGACGGAGGACACTAGCCAGCCCGGAC	1144
OY	1365	CGCGTGCACAGCTTCACTTTCCCTGAGACTCAGTTTATGCGCGTCAACCGCTTACCAAGAAC	1422
Db	1141	CGCGTGCACAGCTTCACTTTCCCTGAGACTCAGTTTATGCGCGTCAACCGCTTACCAAGAAC	1200
OY	1425	ACGGAATTTACACACTGAAAAATAGTACAAACCTTTTGGCAAAAGATTTGGGATTAAT	1488

Db	1201	ACGGATATTACACAACTGAAATATGATCAACACCTTTTGGCAAAAGATTTTGGGATAT	1260
Qy	1485	TATGACAGATCTTACACCGGCTGTGACATGACCGCTTGACCCCTCCGCCAACACTGCG	1544
Db	1261	TATGACACATCTTACACCGGCTGTGACATGACCGCTTGACCCCTCCGCCAACACTGCG	1320
Qy	1545	CCGCGCTGCAGATCGTGTCCCGGGGCGGCTGACATGACCGGCTCTTTCTCTGAGAGAC	1604
Db	1321	CCGCGCTGCAGATCGTGTCCCGGGGCGGCTGACATGACCGGCTCTTTCTCTGAGAGAC	1380
Qy	1605	CAGTTCTGTAGACACTACGCCAAGGCGGCTTTCACCCGCGGCGCGGCGCGCGCGCG	1664
Db	1381	CAGTTCTGTAGACACTACGCCAAGGCGGCTTTCACCCGCGGCGCGGCGCGCGCGCG	1440
Qy	1665	CCGCGTACCGAGCCGCGAGCTGCGCGACACACACCGGCTGTCTGCGCGGAGAGAGCCGAG	1724
Db	1441	CCGCGTACCGAGCCGCGAGCTGCGCGACACACACCGGCTGTCTGCGCGGAGAGAGCCGAG	1500
Qy	1725	GACCCGCGGCGCGCGCTGCGCGGACGCTGTTTGTGACGCGCGGCGCAACCGGCTGAGC	1784
Db	1501	GACCCGCGGCGCGCGCTGCGCGGACGCTGTTTGTGACGCGCGGCGCAACCGGCTGAGC	1560
Qy	1785	TTTCGCGGCTCGGCTATGACACGCGCACGACTTTCGCGGCGACGCGGCGACGCTGCTC	1844
Db	1561	TTTCGCGGCTCGGCTATGACACGCGCACGACTTTCGCGGCGACGCGGCGACGCTGCTC	1620
Qy	1845	TTTTCACGCGGCGCGGCGCTGAAAGCGCTGCGCGCTGACAGGCTGACGCTGACGCGCGC	1904
Db	1621	TTTTCACGCGGCGCGGCGCTGAAAGCGCTGCGCGCTGACAGGCTGACGCTGACGCGCGC	1680
Qy	1905	CCGCTCGGCTACTACGACCGGCGGCTGCGGCTGCGGCGGCGCGGCGGCGGCGGCGGCG	1964
Db	1681	CCGCTCGGCTACTACGACCGGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCG	1740
Qy	1965	GGCACCAGTTCGCGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2017
Db	1741	GGCACCAGTTCGCGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1793
RESULT 7			
BC029289		1793 bp	mRNA linear PRI 31-MAR-2005
LOCUS			
DEFINITION		Human sapiens T-box, brain, 1, mRNA (CDNA clone IMAGE:5181773),	
ACCESSION		BC029289	
VERSION		BC029289.1	GI:20379748
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
AUTHORS		1 (bases 1 to 1793)	
		Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,	
		Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,	
		Altschul, S.F., Zeeberg, B., Buetow, K.H., Schefer, C.F., Bat, N.K.,	
		Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,	
		Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,	
		Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,	
		Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,	
		Carinci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,	
		Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,	
		McKernan, K.J., Malek, J.A., Gunaratne, F.H., Richards, S.,	
		Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,	
		Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,	
		Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,	
		Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,	
		Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,	
		Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,	
		Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallwood, D.E.,	
		Schierch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
CONSRMT		Mammalian Gene Collection Program Team	

Qy	1247932	JOURNAL PUBMED REFERENCE AUTHORS CONSRMT TITLE JOURNAL	1260
Db	1247932	JOURNAL PUBMED REFERENCE AUTHORS CONSRMT TITLE JOURNAL	1320
Qy	1545	CCGCGCTGCAGATCGTGTCCCGGGGCGGCTGACATGACCGGCTCTTTCTCTGAGAGAC	1604
Db	1321	CCGCGCTGCAGATCGTGTCCCGGGGCGGCTGACATGACCGGCTCTTTCTCTGAGAGAC	1380
Qy	1605	CAGTTCTGTAGACACTACGCCAAGGCGGCTTTCACCCGCGGCGCGGCGCGCGCGCG	1664
Db	1381	CAGTTCTGTAGACACTACGCCAAGGCGGCTTTCACCCGCGGCGCGGCGCGCGCGCG	1440
Qy	1665	CCGCGTACCGAGCCGCGAGCTGCGCGACACACACCGGCTGTCTGCGCGGAGAGAGCCGAG	1724
Db	1441	CCGCGTACCGAGCCGCGAGCTGCGCGACACACACCGGCTGTCTGCGCGGAGAGAGCCGAG	1500
Qy	1725	GACCCGCGGCGCGCGCTGCGCGGACGCTGTTTGTGACGCGCGGCGCAACCGGCTGAGC	1784
Db	1501	GACCCGCGGCGCGCGCTGCGCGGACGCTGTTTGTGACGCGCGGCGCAACCGGCTGAGC	1560
Qy	1785	TTTCGCGGCTCGGCTATGACACGCGCACGACTTTCGCGGCGACGCGGCGACGCTGCTC	1844
Db	1561	TTTCGCGGCTCGGCTATGACACGCGCACGACTTTCGCGGCGACGCGGCGACGCTGCTC	1620
Qy	1845	TTTTCACGCGGCGCGGCGCTGAAAGCGCTGCGCGCTGACAGGCTGACGCTGACGCGCGC	1904
Db	1621	TTTTCACGCGGCGCGGCGCTGAAAGCGCTGCGCGCTGACAGGCTGACGCTGACGCGCGC	1680
Qy	1905	CCGCTCGGCTACTACGACCGGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCG	1964
Db	1681	CCGCTCGGCTACTACGACCGGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCG	1740
Qy	1965	GGCACCAGTTCGCGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2017
Db	1741	GGCACCAGTTCGCGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1793
FEATURES			
source		Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
gene		1 (bases 1 to 1793)	
CDS		Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carinci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, F.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallwood, D.E., Schierch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
ORIGIN		Query Match Best local similarity 100.0%; Pred. No. 1,7e-221; Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	225	GTCGTGATTTCTAGATTATGATTAATAGACTTTTAAAAACACGAGACGGAGCGGAG 284	
Db	1	GTCGTGATTTCTAGATTATGATTAATAGACTTTTAAAAACACGAGACGGAGCGGAG 60	


```

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 183666)
Waterston,R.H.
Direct Submission
Submitted (04-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 183666)
Waterston,R.H.
Direct Submission
Submitted (07-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 183666)
Waterston,R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 183666)
Wilson,R.K.
Direct Submission
Submitted (21-APR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Mar 24, 2000 this sequence version replaced gi:6042135.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watsn.wustl.edu
----- Summary Statistics
Center project name: H_NH0334E15
-----

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPc1-11 human BAC library was made from the blood of one male
donor, as described by Osogawa,K., Moon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanesi,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-26B22, 200 base pair
overlap. Actual start of this clone is at base position 166089 of
RP11-26B22; actual end of this clone is at base position 183666 of
RP11-334E15.

FEATURES
Source
location/Qualifiers
1..183666
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-334E15"
/clone_11b="RPc1-11"
12580..38516
gene

```

```

mRNA      /gene="PSPMD14"
           join(12580..12687,18220..18294,22194..22319,35953..36015,
           38418..38516)
           /gene="PSPMD14"
CDS        join(<12580..12687,18220..18294,22194..22319,35953..36015,
           38418..38516)
           /gene="PSPMD14"
           /note="Homo sapiens proteasome (prosome, macropain) 26S
           subunit, non-ATPase, 14 (PSPMD14), mRNA.; H_NH0334E15.1
           This gene was based on gi(5031980)
           Continued from H_NH0026B2.2"
           /codon_start=1
           /product="unknown"
           /protein_id="AAI15016.1"
           /db_xref="GI:62822468"
           /translation="VVIFAFLINANMVLGHEPROTTSNLGNKPSIQALIHGLNR
           HYSTITRNKNELFOKMLNHLIKSKSMGGLTLQDSEHCENESYVKEMLELAKYN
           KAYEEEDKMPEOLAIKVNGKDPPKRHLBEHDVLMITSNIYQCLAMLDIVYEK"
misc_feature 41394..42018
             /note="Cpg_island (%GC=73.5, o/e=0.71, #CpGs=58)"
gene          41325..52178
             /gene="TBRI1"
             join(43225..44218,44792..44946,45317..45438,46008..46166,
             47312..47373,50485..52178)
             /gene="TBRI1"
             join(43527..44218,44792..44946,45317..45438,46008..46166,
             47312..47373,50485..51343)
             /gene="TBRI1"
             /note="Homo sapiens T-box, brain, 1 (TBRI1), mRNA.;
             H_NH0334E15.2
             This gene was based on gi(22547231)"
             /codon_start=1
             /product="unknown"
             /protein_id="AAI15017.1"
             /db_xref="GI:62822469"
             /translation="MQLEHTISPTIMLSKKPLNVSSYPHGSGSELVJLDDPIISTTD
             NERSRSLPKKITRGMTNQSDTDFPDSCKSPEGDVQRSKLSPYLVDIGSELRHSFDSAA
             LRLNSQSOSPQSAATAVPASMPERPGCHGPAPHPAFETGSFSRYMAHPVTIGATNSL
             DSSSPQGYTPAGPYPOQYGHSGYOGAPYQFSSTPGLVPOKAQVLCNRLMKLFH
             RHOTEINIITKQGRMRPFPLSFNISGLDPFAHYNI FVDVILADPNHRFGCGKMWPCGK
             ADTNQGMRYVMHDPSPNTGAHMWROEISFGKLTKNNKGASNNNGOMVLIOSLHKYO
             PRHVAEVNDEDTGETDSOPGRVOTFTPEPTPOLAVYAONTIDLTOKIDHNPAFGAG
             DNVDITYTGCDMRLTPSRPDRSQIVYGACARYMAAGSLTQQOFPANSYAKARPFACG
             AGCGPGTRDVPHPTNGLSLSPQAEDBGASPORFWTFPANRLDPAFASVDTATDPAG
             NAATLLSYAAGVAPLQAGACTGTPLDYVYVSOWGASPPQVCSTGSGSVLPWP
             NSAAAAARMAGNPVLYGEBAEGIAAERSLIPCAEDADAPKDLISSSTWETSSTKSI
             DSDDSGIVEQAKRRRISPADTPVSESSSLKSEVLARDCENCKAKDISGYGVFYSHS
             "
misc_feature 43900..44330
             /gene="TBRI1"
             /note="Cpg_island (%GC=63.3, o/e=0.70, #CpGs=30)"
             45767..46201
             /gene="TBRI1"
             /note="Cpg_island (%GC=67.8, o/e=0.83, #CpGs=45)"
             50441..51502
             /gene="TBRI1"
             /note="Cpg_island (%GC=71.4, o/e=0.85, #CpGs=134)"
             54187..55340
             /note="Cpg_island (%GC=62.6, o/e=0.81, #CpGs=99)"
             91821..92708
             /note="Cpg_island (%GC=60.1, o/e=0.83, #CpGs=20)"
             134790..135005
             /note="Cpg_island (%GC=77.8, o/e=0.97, #CpGs=37)"
             172044..172317
             /note="Cpg_island (%GC=60.6, o/e=0.78, #CpGs=21)"
ORIGIN
Query Match      53.2%; Score 1695.4; DB 8; Length 183666;
Best Local Similarity 99.9%; Pred. No. 3.je-209;
Matches 1696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1490 CACGATTACACGGCGCTGTACATGACACGCCGCTGACCCCTCGCCCAAGATCGGCGGG 1549

```



```

Howell, L. T., Huiyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafatis, K., Kelly, S.,
Keys, T., Khan, Z., King, L., Koyar, C., Kovas, A., Kovis, C., Latta, P.,
Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, K., Liu, Y., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorenshewa, L., Lozado, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McLellan, H., McPherson, J., Mercado, C., Mettler, M.,
Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Bastaghn, E.,
Noct, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Oden, E.,
Okunolu, G., Okunolu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Pitman, E., Pu, L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Umani, K., Vargo, C., Verduzco, D., Villasana, D., Vitek, D.,
Vojkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 101040)
Worley, K.C.
Direct Submission
Submitted (17-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
-----
Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help.tmc.edu
-----
Project Information
Center project name: FYTU
Center clone name: CH240-161A23
-----
Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 98251 bases at least Q40
Consensus quality: 103096 bases at least Q20
Consensus quality: 106131 bases at least Q20
Estimated insert size: 124062; sum-of-coverage estimation
Estimated insert size: 107277; agarose-fp estimation
Quality coverage: 2x in Q20 bases; agarose-fp estimation
Quality coverage: 2x in Q20 bases; sum-of-coverage estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1
4966: contig of 4966 bp in length
4967
5066: gap of unknown length
5067
5681: contig of 2515 bp in length
5682
7681: gap of unknown length
7682
11922: contig of 4241 bp in length
11923
12022: gap of unknown length
12023
16702: contig of 4680 bp in length
16703
16802: gap of unknown length
16803
18881: contig of 2079 bp in length
18882
18981: gap of unknown length
22999: contig of 4018 bp in length

```

FEATURES

```

23000 23099: gap of unknown length
23100 25642: contig of 2543 bp in length
25643 25742: gap of unknown length
25743 31586: contig of 5844 bp in length
31587 31687: gap of unknown length
31688 34315: contig of 2629 bp in length
34316 34415: gap of unknown length
34416 37369: contig of 2954 bp in length
37370 37370: gap of unknown length
37470 39730: contig of 2261 bp in length
39731 39830: gap of unknown length
39831 42816: contig of 2986 bp in length
42817 42916: gap of unknown length
42917 48470: contig of 5554 bp in length
48471 48570: gap of unknown length
48571 51440: contig of 2870 bp in length
51441 51540: gap of unknown length
51541 53825: contig of 2285 bp in length
53826 53925: gap of unknown length
53926 56237: contig of 2312 bp in length
56238 56337: gap of unknown length
56338 59002: contig of 2665 bp in length
59003 59102: gap of unknown length
59103 67716: contig of 8614 bp in length
67717 67816: gap of unknown length
67817 71710: contig of 3894 bp in length
71711 71810: gap of unknown length
71811 74047: gap of 2237 bp in length
74048 74147: gap of unknown length
74148 77776: contig of 3629 bp in length
77777 77876: gap of unknown length
77877 86904: contig of 9028 bp in length
86905 87004: gap of unknown length
87005 89555: contig of 2551 bp in length
89556 89556: gap of unknown length
89557 91999: contig of 2344 bp in length
92000 92099: gap of unknown length
92100 95461: contig of 3362 bp in length
95462 95561: gap of unknown length
95562 97821: contig of 2260 bp in length
97822 97921: gap of unknown length
97922 101040: contig of 3119 bp in length.
Location/Qualifiers
1..101040
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-161A23"
4967..5066
/estimated_length=unknown
7582..7681
/estimated_length=unknown
11923..12022
/estimated_length=unknown
16703..16802
/estimated_length=unknown
18882..18981
/estimated_length=unknown
23000..23099
/estimated_length=unknown
25643..25742
/estimated_length=unknown
31587..31686
/estimated_length=unknown
34316..34415
/estimated_length=unknown
37370..37469
/estimated_length=unknown
39731..39830
/estimated_length=unknown
42817..42916
/estimated_length=unknown
48471..48570

```

gap

Query Match 40.5%; Score 1289; DB 14; Length 101040;
Best Local Similarity 88.0%; Pred. No. 7, DE-157;
Matches 1517; Conservative 0; Mismatches 170; Indels 36; Gaps 9;

gap /estimated_length=unknown
51441..51540
/estimated_length=unknown
53826..53925
/estimated_length=unknown
56238..56337
/estimated_length=unknown
59003..59102
/estimated_length=unknown
67717..67816
/estimated_length=unknown
71711..71810
/estimated_length=unknown
74048..74147
/estimated_length=unknown
77777..77876
/estimated_length=unknown
86905..87004
/estimated_length=unknown
89556..89655
/estimated_length=unknown
92000..92099
/estimated_length=unknown
95462..95561
/estimated_length=unknown

QY 1490 CACGATCTACACCGGCTGTGACATGAGCCGCTGACCCCTGCGCCCAAGACTCGCCGG 1549
DB 68037 CAGGATCTACACCGGCTGTGACATGAGCCGCTGACCCCTGCGCCCAAGACTCGCCGG 68096
QY 1550 CTGCGAGATCTGTCGCCGGGGCCCGCTACGCCCATGCGCGGCTCTTCTGCGAGACAGTT 1609
DB 68097 CTGCGAGATCTGTCGCCGGGGCCCGCTACGCCCATGCGCGGCTCTTCTGCGAGACAGTT 68156
QY 1610 CGTGAAGCACTACGCGCAAGGCGCGCTTTCACCGGGCGGGCGCGGGCGCGGGCGGG 1669
DB 68157 CGTGAAGCACTACGCGCAAGGCGCGCTTTCACCGGGCGGGCGCGGGCGCGGGCGGG 68216
QY 1670 TACGAGCCGAGCGTCCGCAACCAACGAGGCTGTGCGCGAGAGAGAGAGAGAGAGAGAG 1729
DB 68217 CACGAGCCGAGCGTCCGCAACCAACGAGGCTGTGCGCGAGAGAGAGAGAGAGAGAGAG 68276
QY 1730 GGGCGCGCTCGCGCGCAACGCTGTTTGTGACCGCGGCAACACCGGCTGGAATTCCG 1789
DB 68277 GGGCGCGCTCGCGCGCAACGCTGTTTGTGACCGCGGCAACACCGGCTGGAATTCCG 68336
QY 1790 GGGCTGGGCTTGAAGACGCGGCACTTCGCGGGCAACGAGGCGAGCGCTGCTCTTA 1849
DB 68337 GGGCTGGGCTTGAAGACGCGGCACTTCGCGGGCAACGAGGCGAGCGCTGCTCTTA 68396
QY 1850 CCGGCGGCGGCGGTGAAGCGCTGCGCTGAGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1909
DB 68397 CCGGCGGCGGCGGTGAAGCGCTGCGCTGAGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCT 68456
QY 1910 CCGGCTACTAGCGGACCGCTGCGGCTGCGGCGCGCGGCACTCCCGGAGTACTGCGGAC 1969
DB 68457 CCGGCTACTAGCGGACCGCTGCGGCTGCGGCGCGCGGCACTCCCGGAGTACTGCGGAC 68516
QY 1970 CAAGTGGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2029
DB 68517 CAAGTGGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68576
QY 2030 CCGCGCCCAATCTTACTTGGGCGAGAGAGCGCGGCTGAGCGCGCGGAGCGCTGCGCT 2089
DB 68577 CCGCGCCCAATCTTACTTGGGCGAGAGAGCGCGGCTGAGCGCGCGGAGCGCTGCGCT 68636
QY 2090 GCGGCGCGCGCGCGCGAGAGAGCGCAAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2149
DB 68637 GCGGCGCGCGCGCGCGAGAGAGCGCAAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 68696

QY 2150 GAGCGCTCTCTGATCAATGCTCATTCAGCTCCAGGAGCTCGGGATTTACGAGAGAGCCAA 2209
DB 68697 GAGCGCTCTCTGATCAATGCTCATTCAGCTCCAGGAGCTCGGGATTTACGAGAGAGCCAA 68756
QY 2210 GCGAGGCGGATCTCGCGGCGGCAACGCGCGTGTCCGAGAGTTGTCCTCCGCTCAAGAG 2269
DB 68757 ACAGAGGCGGATCTCGCGGCGGCAACGCGCGTGTCCGAGAGCTGTCCTCCGCTCAAGAG 68816
QY 2270 CGAGTCTGAGCCAGCGGAGCTGCGGAGAGAACTGCGGCAAGAGACTTGGCGGCTACTA 2329
DB 68817 CGAGTCTGAGCCAGCGGAGCTGCGGAGAGAACTGCGGCAAGAGACTTGGCGGCTACTA 68876
QY 2330 TGGCTTACTTCGACAGGCTAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2386
DB 68877 CGGCTTACTTCGACAGGCTAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 68936
QY 2387 ---ACCCAGCGAGCCCTCAGCTCTTCCCGAGCTCGGCTCCGCACTCTCTCTT 2443
DB 68937 CCGGCGCGCTCGCGGCTCTCGCGCTCTTCCCGAGCTCTGCTCAGCGCCACCTC 68996
QY 2444 GCGCAACCACTATTTATTTGACCTCTGATGCGCTGTGAGAGAAATGAGAGAGTCT 2503
DB 68997 GCGCACTCTCTCAGTTATTTGACCTGCAACACCGCTGAGAGAAATGAGAGAGTCT 69056
QY 2504 CC-GAGGCTATTTAAC-----CTTTTTCAGAGAGCTCTGCAATTA 2548
DB 69057 CCAAGTGTGATTTTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 69116
QY 2549 GCTACCGAGCTTCAACTTGTCTGTAAC--TTTGTGTTTCTTACTTCTTCTTCTG 2607
DB 69117 GCGACCGAGCTTCAACTTGTCTGTAACCTTTTGTGTTTCTTACTTCTTCTG 69176
QY 2608 TGAAGTATCTT-CTTACATTTCCCTTCCCTGCTCTTCTTCTTCTTCTTCTTCTTCT 2666
DB 69177 TGAAGTATCTTACATTTCCCTTCCCTGCTCTTCTTCTTCTTCTTCTTCTTCT 69236
QY 2667 TTTCTTGAATGAATCTTTCACCTTTTGAAGAGAGCGGAGTCTGAGGAGAGAGAGAG 2726
DB 69237 TTTCTATATGAATCTTTCACCTTTTGAAGAGAGCGGAGTCTGAGGAGAGAGAGAG 69296
QY 2727 TTTCCAGCCGCGCAAGTCTCGGCTCTCCACATTAACATAGATGTTGACTTGAACCTGG 2786
DB 69297 TTTCCAGCCGCGCAAGTCTCGGCTCTCCACATTAACATAGATGTTGACTTGAACCTGG 69356
QY 2787 ACCCAACGAGCGCTCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATGATG 2846
DB 69357 ACCCAACGAGCGCTCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATG 69412
QY 2847 ATGTTAATTAATTTTGTGAACAAGCGTGTGAATGATGATGATGATGATGATGATGAT 2906
DB 69413 ATGTTAATTAATTTTGTGAACAAGCGTGTGAATGATGATGATGATGATGATGAT 69472
QY 2907 GTAAGAAATGCTATTTTATTC---TGTCAAGGCAAAACAGTTCATGCTTAAAC 2963
DB 69473 GTAAGAAATGCTATTTTATTCATTTGTCAAGGCAAAACAGTTCATGCTTAAAC 69532
QY 2964 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3023
DB 69533 TTTCT---CATTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 69586
QY 3024 TTTTAAATTTCTTGTGAAGTATTAATTTGAAGGCTTGAAGAAATGATGATGATGATG 3083
DB 69587 TTTTCTTGTGAAGTATTAACATTAATTTGAAGGCTTGAAGAAATGATGATGATGATG 69646
QY 3084 TGAATGGTTTCCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3143
DB 69647 TGAATGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 69706
QY 3144 CACACAAATAGCTAAGAGAGATCAACCAACCTTTAAGG 3186
DB 69707 CACACAAATAGCTAAGAGAGATCAACCAACCTTTAAGG 69749

RESULT 10
AL845291
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-1011 on chromosome 2, complete sequence.
ACCESSION AL845291
VERSION AL845291.5 GI:23954047
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurionathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 209914)
REFERENCE
AUTHORS Blakey,S.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Oct 13, 2002 this sequence version replaced gi:23954047.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (1.e., phred quality >= 30) ; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-1011 is from the RPI-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6.

FEATURES
source
1.209914
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-1011"
/clone_lib="RPI-23"

ORIGIN
Query Match 36.6%; Score 1164.6; DB 9; Length 209914;
Best Local Similarity 85.6%; Pred. No. 7.1e-141;
Matches 1494; Conservative 0; Mismatches 179; Indels 72; Gaps 15;

1490 CACGATCTACCGGCTGTGACATGACCGCTGACCGCCCTGCCCAAGACTGCGCGG 1549
|||
Db 186044 CAGGATCTACACGGGCTGTGACATGACCGCTGACCGCCGTGCCCAAGACTCTCCGG 186103
|||
Qy 1550 CTCGAGATCTGCGCGCGCGCGCGCTAGCGCATGCGCGGCTTTCTCGAGAGCAAGTT 1609
|||
Db 186104 CTCGAGATCTGCGCGCGCGCGCTAGCGCATGCGCGGCTTTCTCGAGAGCAAGTT 186163
|||
Qy 1610 CGTGAGCACTACGCGCAAGCGCGCTTCACCGCGCGCGCGCGCGCGCGCGCGCG 1669
|||

Db 186164 CGTGAGCACTACGCGCAAGCGCGCTTCACCGCGCGCGCGCGCGCGCGCGCGCG 186223
Qy 1670 TACGAGACCGGAGCGTGGCCGACACCAACGGGCTGTCTCCCGGAGAGGCCAGAGACC 1729
|||
Db 186224 CACGAGACCGGAGCGTGGCCGACACCAACGGGCTGTCTCCCGGAGAGGCCAGAGACC 186283
|||
Qy 1730 GGGCGCGCGCGCGCGGACGCTGGTTGTGACGCGCGGACCAACCGGCTGACATTGCG 1789
|||
Db 186284 GGGCGCGCGCGCGCGGACGCTGGTTGTGACGCGCGGACCAACCGGCTGACATTGCG 186343
|||
Qy 1790 GGGCTGAGCTATGACACGCGCACGACCTTGCGGGGACGCGGCGACGCTGCTCTTGA 1849
|||
Db 186344 GGGCTGAGCTATGACACGCGCACGACCTTGCGGGGACGCGGCGACGCTGCTGCTGA 186403
|||
Qy 1850 CGCGCGCGCGCGCGGCTGGAAGCGCTGCGGCTGACGAGGCTGACATGCGCGCGCGCT 1909
|||
Db 186404 TGCAGCGCGCGCGCTGGAAGCGCTGCGGCTGACGAGGCTGACATGCGCGCGCGCT 186463
|||
Qy 1910 CGGCTACTACGCGCACCGCGCTGCGGCGCGCGCGAGTCCCGGAGTACTGCGGAC 1969
|||
Db 186464 CGGCTACTACGCGCACCGCGCTGCGGCGCGCGAGTCCCGGAGTACTGCGGCGC 186523
|||
Qy 1970 CAAGTCGAGCTCGGTCGTGCTGCTGCTGAGCCCAACAGCGCGCGCGCGCGCATGCG 2029
|||
Db 186524 CAAGTCGAGCTCGGTCGTGCTGCTGAGCCCAACAGCGCGCGCGCGCGCATGCG 186583
|||
Qy 2030 CGCGCGCAATCCCTACTCTGGCGGAGAGCGCGGCGCTGCGCGCGCGCGCGCTGCGCGCT 2089
|||
Db 186584 CGCGCGCAATCCCTACTCTGGCGGAGAGCGCGGCGCTGCGCGCGCGCGCGCTGCGCGCT 186643
|||
Qy 2090 GCCCGCGCGCGCGCGGAGAGCGCGCAAGGCCAAGGACCTGTCGANTTCAGCTGATGCA 2149
|||
Db 186644 GCGCGC---CGCGCGGAGAGCGCGCAAGGCCAAGGACCTGTCGANTTCAGCTGATGCA 186700
|||
Qy 2150 GAGCGCTCTCTGATCAAGTCATGCACTCCAGGACTCGGGATTTACAGAGCGCAA 2209
|||
Db 186701 GAGCGCTCTCTGATCAAGTCATGCACTCCAGGACTCGGGATTTACAGAGCGCAA 186760
|||
Qy 2210 GCGAGGCGGATCTGCGCGCGCGCGGACGCGCGCTGCGGAGATTCGTCGCGCGCAAGAG 2269
|||
Db 186761 GCGAGGCGGATCTGCGCGCGCGCGGACGCGCGCTGCGGAGATTCGTCGCGCGCAAGAG 186820
|||
Qy 2270 CGAGTCTGCGCGCGGAGGAGCTGCGAGAAAGTGGCGCAAGGACTTATGAGCGCTACTA 2329
|||
Db 186821 CGAGTCTGCGCGCGGAGGAGCTGCGAGAAAGTGGCGCAAGGACTTATGAGCGCTACTA 186880
|||
Qy 2330 TGGCTTCTACTGCGACAGCTATGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGGAC 2389
|||
Db 186881 TGGCTTCTACTGCGACAGCTATGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGGAC 186940
|||
Qy 2390 CCCAGCGAGCGCGCTGCGAGCGCTTCCCGAGCTCGCGCTCCCGACATCTCTGCGGCGAC 2449
|||
Db 186941 CC-----CCAGCGTGGCGCGCGCGCGCGCTCTCTCCCGAGCGCGCTTCTAGCGCAC 186988
|||
Qy 2450 CCACTCATTTTATTTGACCGCTCGATGCGCGCTGCGACGATATAGTCAAGTCTGCGAGC 2509
|||
Db 186988 TCGGCTC-TTTCACCTTGAACCTCGATGACGCTGCG-GGGATATAGTCAAGTCTGCACT 187046
|||
Qy 2510 GTGATTTTAACTTTT-----TGACAGCAAGTCTTGCAG 2545
|||
Db 187047 ATGATTTTAAACTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTGAG 187106
|||
Qy 2546 TTACTCATCGGACCTTGAACCTTGGCTTGAACCTTTGGTTTCTTCTTCTTCTTCTTCTT 2605
|||
Db 187107 TTACGCGACGACCTTGAACCTTGGCTTGAACCTTTGGTTTCTTCTTCTTCTTCTTCTT 187162
|||
Qy 2606 TGTGAGTTATCTCTTAC-----AATCCCGCTCCCGCTGCTTCTTCTTCTTCTTCTTCT 2659
|||
Db 187163 TGTGAGTTATGATCT 187222
|||
Qy 2660 CTTCCTTTCTTGTATGAATCTTTCACCTTTTGAAGACCTGAGGACCTCTGTGAGCA 2719
|||
Db 187223 CTTCCTTTCTTGTATGAATCTTTCACCTTTTGAAGACCTGAGGACCTCTGTGAGCA 187282
|||

Oy	3075	ACTCAGTAGGATGGGTTTCCCACTTCTCCTCAATCCGTTGATGAAATAATTAATAGT	3134
Db	86799	AACACTTCATGATGGTGCTTTTTCATCTTCTCAGTTCGTTGATGATGATACCACTAATG	86740
Oy	3135	GGCCTATAGCACAACAATAGCTAAGAGAAATCCACCAAAACACCTTTAAAG	3186
Db	86739	GCCCAAGTGCACACAGCATAGCTTAGAGAAATCCACTGAACTATATATAAG	86687

RESULT 12				
AC165777/c				
LOCUS				
DEFINITION	AC165777	101040 bp	DNA	linear HTG 17-JUL-2005
ACCESION	BOB taarus clone	CH240-161A23, ***	SEQUENCING IN PROGRESS ***	27
	unordered pieces.			
	AC165777			

ACCESSION	AC165777
VERSION	AC165777.1
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Bos taurus (cow)

ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 101040)

Munzy, D., Adams, C., Agbai II, O., Allen, C., Alebrooks, S., Archer, P., Artedomo, H., Bandaratnaie, D., Bangura, L., Beltrán, B., Beltrán, R., Berzaducci, A., Biswal, K., Blyth, P., Bonham, H., Bunay, C., Butch, P., Cadoree, I., Chandra, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chattroum, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Kocha, S., Dunn, A., Durbin, K., Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, A., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gastrik, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, W., Haaland, W., Haeblerlen, K., Hall, B., Hamdt, H., Hamilton, K., Harbes, B., Harris, R., Haylak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Herrands, D., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L.T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Kelly, S., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafits, K., Kelly, S., Keys, T., Khan, Z., King, L., Koyar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., Legall, F.I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Limehl, M., Liu, W., Liu, Y.-S., Liu, Y., LiYanage, D., London, P., Lopez, J., Lorensuewa, L., Lozada, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McClelland, H., McPherson, J., Mercado, C., Metker, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Mundasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Bataugh, E., Nott, A., Nwokoletemeh, O., Obregon, M., Ochi-Okorite, C., Oden, E., Okunou, G., Okunou, K., Parker, D., Pasternack, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puzao, M., Qin, X., Quinn, A., Quiroz, J., Rabara, D., Raehlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojale, A., Ruiz, S.-J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Saverly, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergrén, E., Song, X.-Z., Sorelle, R., Swasek, A., Taylor, E., Sodergrén, T., Thomas, N., Thörn, R., Thornton, R., Trejos, Z., Usmani, K., Varco, C., Varco, C., Verdusco, D., Villasana, D., Virk, D., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Yang, Z., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 101040)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (17-JUL-2005)
	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

 Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help.tmc.edu

 Project Information
 Center project name: FUJU
 Center clone name: CH240-161A23

 Summary Statistics
 Sequencing vector: Plasmid,
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 98251 bases at least Q40
 Consensus quality: 103096 bases at least Q30
 Consensus quality: 106131 bases at least Q20
 Estimated insert size: 124062; sum-of-contigs estimation
 Estimated insert size: 107277; agarose-fp estimation
 Quality coverage: 2x in Q20 bases; agarose-fp estimation
 Quality coverage: 2x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 27 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

*	1	4366:	contig of 4366 bp in length
*	4967	5066:	gap of unknown length
*	5067	7581:	contig of 2515 bp in length
*	7582	7681:	gap of unknown length
*	7682	11922:	contig of 4241 bp in length
*	11923	11202:	gap of unknown length
*	12023	16702:	contig of 4680 bp in length
*	16703	16802:	gap of unknown length
*	16803	18881:	contig of 2079 bp in length
*	18882	18981:	gap of unknown length
*	18982	22999:	contig of 4018 bp in length
*	23000	23099:	gap of unknown length
*	23100	25542:	contig of 2543 bp in length
*	25543	25742:	gap of unknown length
*	25743	31586:	contig of 5844 bp in length
*	31587	31686:	gap of unknown length
*	31687	34315:	contig of 2629 bp in length
*	34316	34415:	gap of unknown length
*	34416	37469:	contig of 2954 bp in length
*	37470	37469:	gap of unknown length
*	37470	39730:	contig of 2261 bp in length
*	39731	39830:	gap of unknown length
*	39831	42816:	contig of 2966 bp in length
*	42817	42916:	gap of unknown length
*	42917	48470:	contig of 5554 bp in length
*	48471	48570:	gap of unknown length
*	48571	51440:	contig of 2870 bp in length
*	51441	51540:	gap of unknown length
*	51541	53825:	contig of 2225 bp in length
*	53826	53925:	gap of unknown length
*	53926	56237:	contig of 2312 bp in length
*	56238	56337:	gap of unknown length
*	56338	59002:	contig of 2665 bp in length
*	59003	59102:	gap of unknown length
*	59103	67716:	contig of 8614 bp in length
*	67717	67816:	gap of unknown length
*	67817	71710:	contig of 3894 bp in length
*	71711	71810:	gap of unknown length
*	71811	74047:	contig of 2237 bp in length
*	74048	74147:	gap of unknown length
*	74148	77776:	contig of 3629 bp in length
*	77777	77876:	gap of unknown length
*	77877	86904:	contig of 9028 bp in length

```
* 86905 87004: gap of unknown length
* 87005 89555: contig of 2551 bp in length
* 89556 89555: gap of unknown length
* 89556 91959: contig of 2344 bp in length
* 92000 92099: gap of unknown length
* 92100 95461: contig of 3362 bp in length
* 95462 95561: gap of unknown length
* 95562 97821: contig of 2260 bp in length
* 97822 97921: gap of unknown length
* 97922 101040: contig of 3119 bp in length.
Location/Qualifiers
1..101040
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-161A23"
4967..5066
/estimated_length=unknown
gap
/estimated_length=unknown
7582..7681
/estimated_length=unknown
11923..12022
/estimated_length=unknown
16703..16802
/estimated_length=unknown
18882..18981
/estimated_length=unknown
23000..23099
/estimated_length=unknown
25643..25742
/estimated_length=unknown
31587..31686
/estimated_length=unknown
34316..34415
/estimated_length=unknown
37370..37469
/estimated_length=unknown
39731..39830
/estimated_length=unknown
42817..42916
/estimated_length=unknown
48471..48570
/estimated_length=unknown
51441..51540
/estimated_length=unknown
53826..53925
/estimated_length=unknown
56238..56337
/estimated_length=unknown
59003..59102
/estimated_length=unknown
67717..67816
/estimated_length=unknown
71711..71810
/estimated_length=unknown
74048..74147
/estimated_length=unknown
77777..77876
/estimated_length=unknown
86905..87004
/estimated_length=unknown
89556..89555
/estimated_length=unknown
92000..92099
/estimated_length=unknown
95462..95561
/estimated_length=unknown
gap
/estimated_length=unknown
Query Match 30.5%; Score 971; DB 14; Length 101040;
Best Local Similarity 98.5%; Pred. No. 6,7e-116;
Matches 980; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
Dh 65426 CAGGTGATTAATCTAATTAATGTCATCTAATTAATTAATGTCAGACGTTAAACCAATGG 65367
Qy 61 CAGAGCGGTTTATTCGGCTGCAACAGACCAAGATCAAAAGGAGCTTTCTGATGTC 120
Dh 65366 CAGAGCGGTTTATTCGGCTGCAACAGACCAAGATCAAAAGGAGCTTTCTGATGTC 65307
Qy 121 TGCATATGTCATATTTGGCCATCTCTTCTCCAGGAGAAAAAGTAATCAAACTTT 180
Dh 65306 TGCATATGTCATATTTGGCCATCTCTTCTCCAGGAGAAAAAGTAATCAAACTTT 65247
Qy 181 TGAGAGCAATTTGCTGGTTGAAGTCTTTCTGCTAGTGAAGGGGCTCTGGAGTTTCTAG 240
Dh 65246 TGAGAGCAATTTGCTGGTTGAAGTCTTTCTGCTAGTGAAGGGGCTCTGGAGTTTCTAG 65187
Qy 241 TTTATGTAATATAGGACTTTAAAAACAGGAGACGAGGAGGAGTTCAGTTCTAGAG 300
Dh 65186 TTTATGTAATATAGGACTTTAAAAACAGGAGACGAGGAGGAGTTCAGTTCTAGAG 65127
Qy 301 CTATGACGCTGAGACCTGCTTTCTCTCTATCATGCTCTCCAAAGAAATTTCTCAATG 360
Dh 65126 CTATGACGCTGAGACCTGCTTTCTCTCTATCATGCTCTCCAAAGAAATTTCTCAATG 65067
Qy 361 TGAGCAGCAGTACCACTTCAAGCGGATCCAGACTTGTTCGACGATCATCCATT 420
Dh 65066 TGAGCAGCAGTACCACTTCAAGCGGATCCAGACTTGTTCGACGATCATCCATT 65007
Qy 421 TCTGACACCTGACCAACTGAGAGAGAGTCACTTTGAAAAAATTTACAGAGGGAGTGA 480
Dh 65006 TCTGACACCTGACCAACTGAGAGAGAGTCACTTTGAAAAAATTTACAGAGGGAGTGA 64947
Qy 481 CGAATCAGTCAATACAGACAAATTTCTGACTTCAAGACTCAACAGGAGGAGTCCAGA 540
Dh 64946 CGAATCAGTCAATACAGACAAATTTCTGACTTCAAGACTCAACAGGAGGAGTCCAGA 64887
Qy 541 GAAGTAACTCTCTCTCTGTTGAGCGGGTCTTGAGCTTCTGACATTTGATGGCT 600
Dh 64886 GAAGTAACTCTCTCTCTGTTGAGCGGGTCTTGAGCTTCTGACATTTGATGGCT 64827
Qy 601 CTGCTGAGATCGGCTACCTCTCTGATCGAGCCAGCCAGTCCGGGCACTGCTC 660
Dh 64826 CTGCTGAGATCGGCTACCTCTCTGATCGAGCCAGCCAGTCCGGGCACTGCTC 64767
Qy 661 CCAATGCCATGTTCCGTTACCCCGCCAGCAGGACCGGCGCAACCCGCTTCTCAATG 720
Dh 64766 CCAATGCCATGTTCCGTTACCCCGCCAGCAGGACCGGCGCAACCCGCTTCTCAATG 64707
Qy 721 GCAGCCCTAGCCGCTATAGGCCCAACCCGCTTATCAACCAAGAGCTTCAACAGCC 780
Dh 64706 GCAGCCCTAGCCGCTATAGGCCCAACCCGCTTATCAACCAAGAGCTTCAACAGCC 64647
Qy 781 TCCGTGCAACTCTGCGCGAGGGATACCCACGCGCGGCTACCCCTACCAACAGAGT 840
Dh 64646 TCCGTGCAACTCTGCGCGAGGGATACCCACGCGCGGCTACCCCTACCAACAGAGT 64587
Qy 841 ACGGCCACTCTTCAACAGAGCTCCGTTTCAACAGTTCTCTCCACCCAGCCGGGGCTGG 900
Dh 64586 ACGGCCACTCTTCAACAGAGAGCCCGCTTCAACAGTTCTCTCCACCAACCCGGGGCTGG 64527
Qy 901 TGCCCGCAAGCAACAGGTGATCTGTGCAACAGGCCCTTTGGCTGAATTTCAACGGC 960
Dh 64526 TGCCCGCAAGCAACAGGTGATCTGTGCAACAGGCCCTTTGGCTGAATTTCAACGGC 64467
Qy 961 ACCAAACGAGATGATCATCAACCAACAGGAAAGG 995
Dh 64466 ACCAAACGAGATGATCATCAACCAACAGGAAAGG 64432
```

RESULT 13
AF287006 2218 bp mRNA linear VRT 10--JAN-2001
LOCUS AF287006
DEFINITION Danilo rerio T-box brain 1 mRNA, partial cds.
ACCESSION AF287006
VERSION AF287006.1 GI:12060813

AUTHORS Plass, C.
TITLE Detection of methylated CpG rich sequences diagnostic for malignant cells
JOURNAL Patent: US 6893820-A 37 17-MAY-2005;
The Ohio State University Research Foundation; Columbus, OH
FEATURES
source
1..514
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 16.1%; Score 514; DB 6; Length 514;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1504 GCTGTGACATGAGACCGCTGACCCCTGCGCCCAACGACTCGCGCGCTGCGAGATCGTGC 1563
DB 514 GCTGTGACATGAGACCGCTGACCCCTGCGCCCAACGACTCGCGCGCTGCGAGATCGTGC 455
QY 1564 CCGGGGGCCCGCTGACCGCATGCGCGGCTCTTTCGACAGACCAAGTTGCTGAGCAACTAG 1623
DB 454 CCGGGGGCCCGCTGACCGCATGCGCGGCTCTTTCGACAGACCAAGTTGCTGAGCAACTAG 395
QY 1624 CCAGAGCCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGGTAACGACGCAACG 1683
DB 394 CCAGAGCCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGGTAACGACGCAACG 335
QY 1684 TGCGGACACCAACGAGCGCTGCTGTGCGCGAGACGCGGAGAGACCGCGCGCGCTTGC 1743
DB 334 TGCGGACACCAACGAGCGCTGCTGTGCGCGAGACGCGGAGAGACCGCGCGCGCTTGC 275
QY 1744 CGCAACGCGTGTGTAACGCGCGCAACACCGGCTGGAATTGCGCGCGCTGCGCTATG 1803
DB 274 CGCAACGCGTGTGTAACGCGCGCAACACCGGCTGGAATTGCGCGCGCTGCGCTATG 215
QY 1804 ACACGCGCACGAGACTTTCGCGGCAACGCGCGCGCGCGCTGCTCTTACGCGCGCGCG 1863
DB 214 ACACGCGCACGAGACTTTCGCGGCAACGCGCGCGCGCGCTGCTCTTACGCGCGCGCG 155
QY 1864 TGAAGCGCTGCGCTGCGAGGCTGAGAGCTGACACTGCGCGCGCTGCGCTAATGCGCG 1923
DB 154 TGAAGCGCTGCGCTGCGAGGCTGAGAGCTGACACTGCGCGCGCTGCGCTAATGCGCG 95
QY 1924 ACCCGTCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1983
DB 94 ACCCGTCGCGGCTGCG 35
QY 1984 TGCTGCTGCTGCG 2017
DB 34 TGCTGCTGCTGCG 1

RESULT 15
HSA341671 793 bp DNA linear PRI 18-JUL-2002
LOCUS HSA341671
DEFINITION Homo sapiens genomic sequence surrounding Not1 site, clone
N16-LP8R.
VERSION AJ341671
ACCESSION AJ341671.1 GI:15886089
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Kuzenko, A.S., Gzatsullin, R.Z., Al-Amin, A.N., Wang, F., Krashinsky, S.M.,
1 (bases 1 to 793)
Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,
Levitsky, V.G., Kolchanov, N.A., Protopopov, A.I., Kashuba, V.I.,
Kisilev, L.L., Wasserman, W., Wahlestedt, C., and Zadorovskiy, E.R.
TITLE Not1 flanking sequences: a tool for gene discovery and verification
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)

PUBMED 12136098
REFERENCE 2 (bases 1 to 793)
AUTHORS Zadorovskiy, E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
source
1..793
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="N16-LP8R"

ORIGIN

Query Match 15.4%; Score 491.2; DB 8; Length 793;
Best Local Similarity 87.0%; Pred. No. 8.2e-54;
Matches 574; Conservative 0; Mismatches 80; Indels 6; Gaps 3;

QY 2010 GCGGCGCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2069
DB 1 GCGGCGCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 2070 GCGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2129
DB 61 GCGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 2130 TCGGATTCGACGCTGATGAGAGCGCCCTCTCGATCAAGTCATCGACTCGAGACTCG 2189
DB 121 TCGGATTCGACGCTGATGAGAGCGCCCTCTCGATCAAGTCATCGACTCGAGACTCG 180
QY 2190 GGAATTTACGAGCGAGCGCAAGCGAGCGAGCTGCGCGCGCGCGCGCGCGCGCGCG 2249
DB 181 GGAATTTACGAGCGAGCGCAAGCGAGCGAGCTGCGCGCGCGCGCGCGCGCGCGCG 240
QY 2250 AGTTGCTCCCGCTCAAGAGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2309
DB 241 AGTTGCTCCCGCTCAAGAGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 2310 AAGGACATTAGCGGCTACTGATGCTTACTGCGACAGCTAGGCGCGCGCGCGCGCG 2369
DB 301 AAGGACATTAGCGGCTACTGATGCTTACTGCGACAGCTAGGCGCGCGCGCGCGCG 360
QY 2370 GCCCG 2429
DB 361 GCCCG 420
QY 2430 CCAACCTCTCTTTCG 2489
DB 421 CCAACCTCTCTTTCG 480
QY 2490 ATTAGTGCAGGCTGCGCGAG-CGTGATTTTAACTTTTTCGACAGAGCTGCGACTTA 2548
DB 481 ATTAGTGCAGGCTTCCAGCGCGGTGATATACCTTTTTCGAGAGGCTGCG-ATAA 539
QY 2549 GCTACCGGACCTTCACTTGTGTAAACCTTTTGTGTTTCTTACTTCTTCTTGTGT 2608
DB 540 GGTACACGCGCTTAATTTTTCGGAACCATTTGGGTTCTC-----TTACCTCTTCTT 595
QY 2609 GGAATTATCTCTTCAATTTCCCTCCCTGCTTCTTCTTCACTCTTCTTCTTCTT 2668
DB 596 GGAAGAGATATCTCCCAAAACCGCGCGCGGTTTATTTCTTCTTCTTCTTCTT 655

Search completed: December 20, 2005, 12:35:54
Job time: 15693 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 03:16:07 ; Search time 1805 Seconds
(without alignments)
11763.829 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186
Sequence: 1 caggtgattatcctaataa.....caccacaacaccttaag 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3186	100.0	3186	12	ADL97817 Human TBR
2	1793	56.3	1793	11	AEA79130 Human apo
3	1695.4	53.2	6878	5	ABA14582 Human ner
4	864	27.1	864	12	ACH87502 Human gen
5	604	19.0	1026	6	ABQ22410 Oligonuc1
6	604	19.0	1026	6	ABQ22411 Oligonuc1
7	515	16.2	515	12	ACH73791 Human gen
8	514	16.1	514	6	ABT07220 Human Cpg
9	469.6	14.7	1026	6	ABQ22413 Oligonuc1
10	469.6	14.7	1026	6	ABQ22412 Oligonuc1
11	419	13.2	457	9	ACH15279 Human adu
12	384.6	12.1	388	3	AAO0887 Human sec
13	374.6	11.8	2494	3	AAZ46477 Human T-b
14	374.6	11.8	3465	12	ADP28768 Human sec
15	363	11.4	389	2	AAQ39695 Expressed
16	363	11.4	389	2	AAQ59107 Human bra
17	357	11.2	2425	11	ADL22644 Human dis
18	350.8	11.0	2883	3	AAZ90784 Mouse ner
19	346	10.9	2061	3	AAZ90785 Human ner

20	346	10.9	2061	12	ADQ19927
21	342.8	10.8	3137	12	ADQ24019
22	341.2	10.7	2756	14	ADY20406
23	341.2	10.7	2756	14	ADY19051
24	321	10.1	463	13	ACF88226
25	316.2	9.9	1608	4	AAQ91988
26	316.2	9.9	1608	9	ADA02650
27	316.2	9.9	1608	9	ACC84950
28	316.2	9.9	1608	10	ADB72388
29	316.2	9.9	1608	10	ADB72388
30	316.2	9.9	1608	10	ADB72388
31	316.2	9.9	2450	14	ADZ39788
32	316.2	9.9	2589	14	ADZ39787
33	316.2	9.9	2589	9	ADA02649
34	316.2	9.9	2589	10	ADB72387
35	316.2	9.9	2589	10	ADB72387
36	316.2	9.9	2589	13	ADP55341
37	316.2	9.9	2589	14	ADX01430
38	316.2	9.9	2589	14	ADY17747
39	316.2	9.9	2590	14	ADZ39814
40	298.6	9.4	1593	9	ADA02647
41	298.6	9.4	1593	10	ADB72385
42	298.6	9.4	1593	10	ADB72385
43	298.6	9.4	2482	9	ADA02646
44	298.6	9.4	2482	10	ADB72384
45	298.6	9.4	2482	10	ADE95894

ALIGNMENTS

RESULT 1
ADL97817
ID ADL97817 standard, DNA, 3186 BP.
XX
XX ADL97817;
AC
XX
DT 03-JUN-2004 (first entry)
XX
DE Human TBR1 DNA SeqID 3.
XX
XX human; gene; ds; neuropsychiatric disorder; bipolar disorder;
XX schizoprenia; alpha-type II calcium/calmodulin dependent protein kinase;
XX CAMK1T-alpha; TBR1; Brachyury; neuroleptic; antidepressant; gene therapy;
XX mental disorder; mood disorder; psychosis; major depression.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 303..2351
FT /ftag= a
FT /product= "TBR1 protein"
XX
XX W02004020455-AZ.
XX
XX 11-MAR-2004.
XX
XX 26-AUG-2003; 2003WO-US026873.
XX
XX 28-AUG-2002; 2002US-0406879P.
XX
XX 27-FEB-2003; 2003US-0451306P.
XX
XX (STRD) UNIV IRLAND STANFORD JUNIOR.
XX
XX Bunney WR, Jones EG, Molnar M;
XX WPI: 2004-239152/22.
XX P-PSDB; ADL97818.
XX
XX Determining whether a subject has or is predisposed to a mental disorder,
XX comprises detecting the level of reagent that selectively associates with
XX a specified polynucleotide or polypeptide in a sample from the subject.
XX

QY 1861 GGTGAAGCGCTGCGCTGCAAGCTGCACTGAGCCGCCGCTCGACTACAG 1920
 DB 1861 GGTGAAGCGCTGCGCTGCAAGCTGCACTGAGCCGCCGCTCGACTACAG 1920
 QY 1921 CGACCCGCTGCG 1980
 DB 1921 CGACCCGCTGCG 1980
 QY 1981 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 DB 1981 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 QY 2041 CCTACCTGCG 2100
 DB 2041 CCTACCTGCG 2100
 QY 2101 CG 2160
 DB 2101 CG 2160
 QY 2161 CGATCAAGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 2220
 DB 2161 CGATCAAGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 2220
 QY 2221 TCTGCG 2280
 DB 2221 TCTGCG 2280
 QY 2281 CCCAGCG 2340
 DB 2281 CCCAGCG 2340
 QY 2341 CGCAGCG 2400
 DB 2341 CGCAGCG 2400
 QY 2401 CCTCAAGCTGCTTCCGAGCTGCTTCCGAGCTGCTTCCGAGCTGCTTCCGAG 2460
 DB 2401 CCTCAAGCTGCTTCCGAGCTGCTTCCGAGCTGCTTCCGAGCTGCTTCCGAG 2460
 QY 2461 ATTTGACCTCGATGCGCTGCTGCGAGCTGCTGCGAGCTGCTGCGAGCTGCT 2520
 DB 2461 ATTTGACCTCGATGCGCTGCTGCGAGCTGCTGCGAGCTGCTGCGAGCTGCT 2520
 QY 2521 CTTTTTGAAGCGAGCTGCTGCGAGCTGCTGCGAGCTGCTGCGAGCTGCTGCG 2580
 DB 2521 CTTTTTGAAGCGAGCTGCTGCGAGCTGCTGCGAGCTGCTGCGAGCTGCTGCG 2580
 QY 2581 TTGGTTTCTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
 DB 2581 TTGGTTTCTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
 QY 2641 GTCCTTCTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700
 DB 2641 GTCCTTCTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700
 QY 2701 TGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
 DB 2701 TGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
 QY 2761 CATAGATGTTGACTCTAGAACCTGAGCCACCGCGCTCTTATCTCCCGAGTG 2820
 DB 2761 CATAGATGTTGACTCTAGAACCTGAGCCACCGCGCTCTTATCTCCCGAGTG 2820
 QY 2821 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2880
 DB 2821 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2880
 QY 2881 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
 DB 2881 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940

QY 2941 ACAAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3000
 DB 2941 ACAAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3000
 QY 3001 TCTCATCTTCT 3060
 DB 3001 TCTCATCTTCT 3060
 QY 3061 TAGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120
 DB 3061 TAGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120
 QY 3121 AATAATTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180
 DB 3121 AATAATTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180
 QY 3181 TAAAGG 3186
 DB 3181 TAAAGG 3186
 RESULT 2
 ID AEA79130 standard; cDNA; 1793 BP.
 XX AEA79130;
 XX 11-AUG-2005 (first entry)
 XX DE Human apoptosis-associated cDNA SEQ ID 378.
 XX XX apoptosis; cell death; cytoskeletal; neuroprotective; immunosuppressive;
 KW antirheumatic; antiarthritic; dermatological; antiinflammatory;
 KW hepatotropic; virucidal; nootropic; anticonvulsant; antiparkinsonian;
 KW vasotropic; cerebroprotective; antialcoholic; gene therapy; tumour;
 KW autoimmune disease; degenerative disease; viral infection; leukaemia;
 KW carcinoma; sarcoma; multiple sclerosis; rheumatoid arthritis; diabetes;
 KW lupus; hepatitis; influenza viruses; Alzheimer's disease;
 KW Huntington's disease; Parkinson's diseases; reperfusion injury; stroke;
 KW alcoholic liver disease; human; gene; ss.
 XX OS Homo sapiens.
 XX XX WO2003058021-A2.
 XX PD 17-JUL-2003.
 XX PF 13-JAN-2003; 2003WO-BP000270.
 XX PR 11-JAN-2002; 2002DE-01000856.
 XX PA (XANT-) XANTOS BIOMEDICINE AG.
 XX PI Koenig-Hoffman K, Kazinski M, Schaefer R, Kesper B;
 DR WPI; 2003-542134/51.
 PT New nucleic acids involved in apoptosis, useful for diagnosis and
 PT treatment of e.g. tumors and degenerative disease, also related proteins,
 PT antibodies and modulators.
 XX PS Claim 1a; SEQ ID NO 378; 517pp; German.
 XX XX This invention describes novel nucleic acid molecules that are associated
 CC with apoptosis and encode a polypeptide and are derived from a normalised
 CC gene library (embryonic or liver) or clone collections, and the extent of
 CC apoptosis measured by cell death detection assay or the CPRG assay
 CC (measuring loss of membrane integrity). The products of the invention
 CC have cytoskeletal, neuroprotective, immunosuppressive, antirheumatic,
 CC antiarthritic, dermatological, antiinflammatory, hepatotropic, virucide,
 CC nootropic, anticonvulsant, antiparkinsonian, vasotropic,
 CC cerebroprotective and antialcoholic activity and can be used for gene
 CC therapy. The polynucleotides also related vectors, hosts (or their

CC extract(s), encoded polypeptide (or their receptors) and/or agents that
CC inhibit their activity (including antisense sequences) are used for
CC treatment or prevention of tumours, autoimmune or degenerative diseases
CC and viral infections, specifically leukaemia, carcinoma, sarcoma,
CC multiple sclerosis, rheumatoid arthritis, diabetes, lupus, or infection
CC with hepatitis or influenza viruses, Alzheimer's, Huntington's or
CC Parkinson's diseases, reperfusion injury, stroke and alcoholic liver
CC disease. Detection of the polymucleotides and derived polypeptides can
CC also be used for diagnosis of these diseases. This sequence encodes an
CC apoptosis-associated protein described in the disclosure of the
CC invention.
CC
XX
XX

Sequence 1793 BP; 385 A; 586 C; 471 G; 351 T; 0 U; 0 Other;

Query Match 56.3%; Score 1793; DB 11; Length 1793;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 225 GTCGTGATTTCTGTTTATGATTAATAGACTTTAAAAACAGAGAGGGGCGAG 284
DB 1 GTCGTGATTTCTAGTTTATATGATTAATAGACTTTAAAAACAGAGAGGGGCGAG 60
OY 285 TGTTCAGGTTCTAGAGCTATGAGCGTGGAGCACTGCTTCTCTTATCATGCTCTCC 344
DB 61 TGTTCAGGTTCTAGAGCTATGAGCGTGGAGCACTGCTTCTCTTATCATGCTCTCC 120
OY 345 AAGAAATTTCTCAATGTGAGCAGCAGCTTACCACATTCAGGGGATCCGAGCTTGTG 404
DB 121 AAGAAATTTCTCAATGTGAGCAGCAGCTTACCACATTCAGGGGATCCGAGCTTGTG 180
OY 405 CACGATCATCCCATATCTCGACCACTGACCACTGGAGAGAGTTCACTTTGAAAAA 464
DB 181 CACGATCATCCCATATCTCGACCACTGACCACTGGAGAGAGTTCACTTTGAAAAA 240
OY 465 ATTACAGAGGGAGTACGAATGATCAGATACAGCAATTTTCTGACTCCAGAGCTCA 524
DB 241 ATTACAGAGGGAGTACGAATGATCAGATACAGCAATTTTCTGACTCCAGAGCTCA 300
OY 525 CCAAGGGAGCTCCAGAGAGTAACTCTCTGCTTTGAGAGGGGCTCTGAGCTTGT 584
DB 301 CCAAGGGAGCTCCAGAGAGTAACTCTCTGCTTTGAGAGGGGCTCTGAGCTTGT 360
OY 585 CACAGTTTCAGATGCTGCTGCTGAGATGCTACTCTCTCTCACTCCAGCGACAG 644
DB 351 CACAGTTTCAGATGCTGCTGCTGAGATGCTACTCTCTCTCACTCCAGCGACAG 420
OY 645 TCTGGGACCACTGCTCCAGTGCATGTTCCCGTAACCCCGGCGAGCAAGCCGCGAC 704
DB 421 TCTGGGACCACTGCTCCAGTGCATGTTCCCGTAACCCCGGCGAGCAAGCCGCGAC 480
OY 705 CCGGCTTCTCATGAGCGAGCCCTAGCCGCTACATGAGCCCAACCCGCTCATCCAC 764
DB 481 CCGGCTTCTCATGAGCGAGCCCTAGCCGCTACATGAGCCCAACCCGCTCATCCAC 540
OY 765 GAGGCTTCAACAGAGCTCTGCTCAACTCTGCGGAGAGGATACCCAGCGGCGCTAC 824
DB 541 GAGGCTTCAACAGAGCTCTGCTCAACTCTGCGGAGAGGATACCCAGCGGCGCTAC 600
OY 825 CCTTACCCACAGAGTACGGCCACTCTACCAAGAGGCTCGTTTACAGTTCTCTCC 884
DB 601 CCTTACCCACAGAGTACGGCCACTCTTACCAAGAGGCTCGTTTACAGTTCTCTCC 660
OY 885 ACCCAGCGGGGCTGTGCTCCGCGCAAGCAAGGTGATCTGTGCAACAGGCGCTTTGG 944
DB 661 ACCCAGCGGGGCTGTGCTCCGCGCAAGCAAGGTGATCTGTGCAACAGGCGCTTTGG 720
OY 945 CTGAATTTTCAACGGACCAAAACGAGATGATCATCAACAAACAGAGAGGCGATGTT 1004
DB 721 CTGAATTTTCAACGGACCAAAACGAGATGATCATCAACAAACAGAGAGGCGATGTT 780
OY 1005 CTTTTTTTAAAGTTTAAATTTCTGCTCGATCCAGGCTCATTAACATATTTTGTG 1064
DB 781 CTTTTTTTAAAGTTTAAATTTCTGCTCGATCCAGGCTCATTAACATATTTTGTG 840

OY 1065 GATGTGATTTTGGGGATATCCCAATCACTGAGGTTTCAAGAGCAATGGTCTTCC 1124
DB 841 GATGTGATTTTGGGGATATCCCAATCACTGAGGTTTCAAGAGCAATGGTCTTCC 900
OY 1125 GGCAGAGCGGACACCAATGTGCAAGAAATCGGCTATATGATCCGATTTCCCAAC 1184
DB 901 GGCAGAGCGGACACCAATGTGCAAGAAATCGGCTATATGATCCGATTTCCCAAC 960
OY 1185 ACTGGGGCTCACTGATGCGCCAAAGAAATCTTTTGAATAATTAACCTTAGACAC 1244
DB 961 ACTGGGGCTCACTGATGCGCCAAAGAAATCTTTTGAATAATTAACCTTAGACAC 1020
OY 1245 AAGAGAGTTCAATTAACATGAGGAGATGGTGTGTTTACAGTCCCTGCAACATACAG 1304
DB 1021 AAGAGAGTTCAATTAACATGAGGAGATGGTGTGTTTACAGTCCCTGCAACATACAG 1080
OY 1305 CCGGCTGATGTGTGGAAGTGAAGAGACGCGACGAGAGCACTAGCCAGCCGCG 1364
DB 1081 CCGGCTGATGTGTGGAAGTGAAGAGACGCGACGAGAGCACTAGCCAGCCGCG 1140
OY 1365 CGGTCGAGACGTTCACTTCTCTGAGACTCATGTTCAATCCGCTACCGCTTACAGAAC 1424
DB 1141 CGGTCGAGACGTTCACTTCTCTGAGACTCATGTTCAATCCGCTACCGCTTACAGAAC 1200
OY 1425 ACGGATATTACAACTGAAAAATGATCAACAACCTTTTGAAGAGATTTGGGATAT 1484
DB 1201 ACGGATATTACAACTGAAAAATGATCAACAACCTTTTGAAGAGATTTGGGATAT 1260
OY 1485 TATGACAGATCTACACCGGCTGTGATGAGACCGCTGACCCCTGCGCCACAGACTCG 1544
DB 1261 TATGACAGATCTACACCGGCTGTGATGAGACCGCTGACCCCTGCGCCACAGACTCG 1320
OY 1545 CCGGCTGAGATGTGTCGCCGGGGCCCGCTACGCCATGCGCGGCTTTCTGACAGAC 1604
DB 1321 CCGGCTGAGATGTGTCGCCGGGGCCCGCTACGCCATGCGCGGCTTTCTGACAGAC 1380
OY 1605 CAGTTTCGAGCACTACAGCCAGAGCCGCTTCCACCCGGGCGGGCGGGCGCCGGG 1664
DB 1381 CAGTTTCGAGCACTACAGCCAGAGCCGCTTCCACCCGGGCGGGCGGGCGCCGGG 1440
OY 1665 CCGGGTACGAGCCGACAGTGCAGCAACCAAGAGGCTGTGTCGCGGACAGGCGGAG 1724
DB 1441 CCGGGTACGAGCCGACAGTGCAGCAACCAAGAGGCTGTGTCGCGGACAGGCGGAG 1500
OY 1725 GACCCGGGCGGCTTCTGCGGCAACGCTGTTGTGACGCGGCGCAACACCGCTGAC 1784
DB 1501 GACCCGGGCGGCTTCTGCGGCAACGCTGTTGTGACGCGGCGCAACACCGCTGAC 1560
OY 1785 TTTGGGGCTTCGGCTTATGACAGGGCCAGAGACTTGGCGGGCAACGCGGCGAGCTG 1844
DB 1561 TTTGGGGCTTCGGCTTATGACAGGGCCAGAGACTTGGCGGGCAACGCGGCGAGCTG 1620
OY 1845 TCTTACCGGCGGCGGCGGCTGTAAGAGCGCTGCGCTGAGAGCTGAGCTGAGCGCG 1904
DB 1621 TCTTACCGGCGGCGGCGGCTGTAAGAGCGCTGCGCTGAGAGCTGAGCTGAGCGCG 1680
OY 1905 CCGCTGAGCTTACAGCCGACCCGCTGGGCTGGGGCGCCCGGAGTCCCGGAGTACTGC 1964
DB 1681 CCGCTGAGCTTACAGCCGACCCGCTGGGCTGGGGCGCCCGGAGTCCCGGAGTACTGC 1740
OY 1965 GGCACCAAGTCCGGCTGTGTGCTGCTGCTGCGCCCAACAGCCCGGCGGCGCG 2017
DB 1741 GGCACCAAGTCCGGCTGTGTGCTGCTGCTGCGCCCAACAGCCCGGCGGCGCG 1793

RESULT 3
ID ABA14582/c
ID ABA14582 standard; DNA; 6878 BP.
XX ABA14582;
XX
DT 23-JAN-2002 (first entry)

PR	1-SEP-2000	2000US -0232899P
PR	14-SEP-2000	2000US -0232400P
PR	14-SEP-2000	2000US -0232401P
PR	14-SEP-2000	2000US -0233063P
PR	14-SEP-2000	2000US -0233064P
PR	14-SEP-2000	2000US -0233065P
PR	21-SEP-2000	2000US -0234223P
PR	21-SEP-2000	2000US -0234747P
PR	25-SEP-2000	2000US -0234977P
PR	25-SEP-2000	2000US -0234988P
PR	26-SEP-2000	2000US -0235848P
PR	27-SEP-2000	2000US -0235834P
PR	27-SEP-2000	2000US -0235836P
PR	29-SEP-2000	2000US -0236377P
PR	29-SEP-2000	2000US -0236368P
PR	29-SEP-2000	2000US -0236369P
PR	29-SEP-2000	2000US -0236370P
PR	29-SEP-2000	2000US -0236680P
PR	02-OCT-2000	2000US -0237037P
PR	02-OCT-2000	2000US -0237038P
PR	02-OCT-2000	2000US -0237039P
PR	02-OCT-2000	2000US -0241186P
PR	20-OCT-2000	2000US -0241187P
PR	20-OCT-2000	2000US -0241808P
PR	20-OCT-2000	2000US -0241809P
PR	20-OCT-2000	2000US -0241826P
PR	20-OCT-2000	2000US -0241827P
PR	20-OCT-2000	2000US -0242211P
PR	01-NOV-2000	2000US -0244674P
PR	08-NOV-2000	2000US -0246477P
PR	08-NOV-2000	2000US -0246478P
PR	08-NOV-2000	2000US -0246532P
PR	08-NOV-2000	2000US -0246533P
PR	08-NOV-2000	2000US -0246534P
PR	08-NOV-2000	2000US -0246552P
PR	08-NOV-2000	2000US -0246556P
PR	08-NOV-2000	2000US -0246577P
PR	08-NOV-2000	2000US -0246582P
PR	08-NOV-2000	2000US -0246532P
PR	08-NOV-2000	2000US -0246609P
PR	08-NOV-2000	2000US -0246610P
PR	08-NOV-2000	2000US -0246611P
PR	08-NOV-2000	2000US -0246613P
PR	17-NOV-2000	2000US -0249207P
PR	17-NOV-2000	2000US -0249208P
PR	17-NOV-2000	2000US -0249210P
PR	17-NOV-2000	2000US -0249211P
PR	17-NOV-2000	2000US -0249212P
PR	17-NOV-2000	2000US -0249213P
PR	17-NOV-2000	2000US -0249214P
PR	17-NOV-2000	2000US -0249215P
PR	17-NOV-2000	2000US -0249216P
PR	17-NOV-2000	2000US -0249217P
PR	17-NOV-2000	2000US -0249218P
PR	17-NOV-2000	2000US -0249244P
PR	17-NOV-2000	2000US -0249245P
PR	17-NOV-2000	2000US -0249246P
PR	17-NOV-2000	2000US -0249255P
PR	17-NOV-2000	2000US -0249257P
PR	17-NOV-2000	2000US -0249259P
PR	17-NOV-2000	2000US -0250391P
PR	01-DEC-2000	2000US -0250391P
PR	01-DEC-2000	2000US -0251030P
PR	05-DEC-2000	2000US -0251988P

Query Match	Best Local Similarity	Matches 1596; Conservative	Score 53.2%;	Pred. No. 0;	Mismatches 0;	Indels 1;	Gaps 0;	Length 6878;	DB 5;	Other:
1490	CACGATCTACACCGGCTGTGATGATGACCGGCTGACCCCTTGCGCCACGACTGCGCGG	1545								
2507	CAGGATCTACACCGGCTGTGATGATGACCGGCTGACCCCTTGCGCCACGACTGCGCGG	2446								
1550	CTCGGAGATCGTGGCCCGGCGGCGCGGCTGACGCGCATGCGCGCTCTTCTTGACGACGAGTT	1605								
2447	CTCGGAGATCGTGGCCCGGCGGCGCGGCTGACGCGCATGCGCGCTCTTCTTGACGACGAGTT	2388								
1610	CGTGAACAATACGCGCAAGGCGCGGCTTCCACCCCGGCGCGGCGCGGCGCGGCGCGG	1666								
2387	CGTGAACAATACGCGCAAGGCGCGGCTTCCACCCCGGCGCGGCGCGGCGCGGCGCGG	2328								
1670	TACGGAACCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1722								
2327	TACGGAACCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2266								
1730	GGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1789								
2267	GGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2205								
1790	GGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1845								
2207	GGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2146								
1850	CGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1905								
2147	CGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2088								

[illegible]


```
DB 721 GAGGGGATCTCGCGCGGACACCCCGGTGTCGAGAGTTGTGTCCCGCTCAAGGCA 780
QY 2273 GGTGCTGGCCCGGAGCTGCGAGAACTGCGCCCAAGACATTAGCGGTACTATGG 2332
DB 781 GGTGTGCGCCCGGAGCTGCGAGAACTGCGCCCAAGACATTAGCGGTACTATGG 840
QY 2333 CTTTACTGCGACAGCTAGGCGCC 2356
DB 841 CTTTACTGCGACAGCTAGGCGCC 864

RESULT 5
ABO22410/C
ID ABO22410 standard; DNA; 1026 BP.
XX
AC ABO22410;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 9001.
XX
KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KM drug; side effect; cancer; central nervous system; cardiovascular;
KM gastrointestinal; respiratory system; single nucleotide polymorphism;
KM SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56bp + Sequence Listing; 56bp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridized to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridization to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridized to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP/s); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABO13410-
CC ABO54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 1026 BP; 140 A; 128 C; 429 G; 329 T; 0 U; 0 Other;
```

```
Query Match 19.0%; Score 604; DB 6; Length 1026;
Best Local Similarity 80.5%; Pred. No. 9.1e-114;
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;

QY 1490 CACGATCTACACGGGCTGTGACATGAGACCGCTTACACCCCTTCGCGCAACGACTGCGCGG 1549
DB 949 CAAATCTACACCGACATATACATTAACCGCTTAAACCCCTTCGCGCAACGACTGCGCGG 890
QY 1550 CTGCGAGATGCTGGCCCGGGGCGCGCTACGACATGAGCGGCTCTTCTTCGAGACGCAATT 1609
DB 889 CTTGCGAANTGCTTACCCGAAACCGGCTACGCCATTAACGACCTCTTCTTCAAAAACCAATT 830
QY 1610 CGTGAGCAACTACGCGCAAGGCGCGCTTCAACCGCGGCGCGGCGCGCGCGCGCGCGCGG 1669
DB 829 CGTAAACAACTACCGCAAAACCGGCTTCAACCGCAAGCGC--GACGCGAACCGCAAGCAA 772
QY 1670 TACGAGCCGACAGCGTGCCTGCGCACACCAACGAGGCTGCTGTGCGGAGAGAGGAGACC 1729
DB 771 TACGAACCGCAACCTTACCGCACACCAACGAACCTACTATCGCGCGCAACCAAGAAACCC 712
QY 1730 GGGGCGCGCCCTGCGCGCAAGCTGTTGTGACGCGCGGCGCAACCGGCTGAGCTTCG 1789
DB 711 GAACGCG--CTGCGCGGACGCTAATTATTAACGCCGACACAAACGACTTAACTTCGCG 653
QY 1790 GGCCTCGGCTTATGACACGCGCAAGCTTCGCGGAGCAAGCGGCAACGCTGCTCTTAA 1849
DB 652 GACCTCGACCTATAACACGACACGAACTTGCGCAACAGCGGACACGCTACTCTCTTA 593
QY 1850 CGCGGCGGCGGCGCTGAAGGCGCTGCGGCTGCAAGGCTGCAAGCTGAGCTGAGCTGCGCGCT 1909
DB 592 CGCGACGACGAACCTTAAACGCT--ACGCTAACAACTAACAACTAACCTAACCGCGCT 534
QY 1910 CGGCTACTACGCGCAACCGCTGCGGCGCGCGGCGCGGAGTCCCGGAGTACTCGGCGAC 1969
DB 533 GCACTACTACGCGCAACCGCTGCGGAGCTTAAACGCGCGCAATCCCGCAATACTACGACAC 474
QY 1970 CAACTCGGAGCTCGGTGCTGCTGCTGCGGCGCAAGCGCGCGCGCGCGCGCATGCG 2029
DB 473 CAAATCGAAGCTCGAAGTCT--CCCTACTAACCAACGACGCGCGCGCGCGCATTAAC 415
QY 2030 CGGCGCCCAATCTCTTACTCTGAGGAGAGAGCGCGAGGCTGCGCGCGCGCGCGCTGCGCGCT 2089
DB 414 CGACGCGCAATCTCTTACTCTTAAACGAAACCACTAACCGCGCGCGCGCGCTGCGCGCT 355
QY 2090 GCCGCGCGGCGCGCGCGCGGAGCGCAAGCGCAAGCGCAAGCGCTGCGGATTCGAGCTGATCGA 2149
DB 354 ACCGCGC--GCGCGCGCGGAAACGCGCAACCGCAAAACCTATCGGATTCGAACTTAATCGA 296
QY 2150 GACGCGCTCTCGATTCAGATCATCGACTCCAGCGACTCGGAGATTACGAGAGCGCAA 2209
DB 295 AACGCGCTCTCGATTCAGATCATCGACTCCAGCGACTCGGAAATTACGAAACAAACGAA 236
QY 2210 GCGGAGCGGATCTCGCGCGCGGAGACAGCGCGGTGTCGAGAGTTGTCCTCCGCTCAAGAG 2269
DB 235 ACGAAACGAATCTCGCGCGGACGCGCGGTATCGGAAATTCGTCCTCCGCTCAAAA 176
QY 2270 CGAGGTGCTGCGCGGAGCGGAGCTGCGAGAACTGCGCGCAAGGACATTAGCGGCTACTA 2329
DB 175 CGAAATVACTTAACCAACGAAACTACGAAAAAATACGAAAAACATTTACGACTACTA 116
QY 2330 TGGCTTACTCTGACAGCTAGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGGAC 2389
DB 115 TAACTTACTCTGACAGCTAAGCGCGCGCTTACCGCGCGCGCGCGCGCGCGCGCGGAC 56
QY 2390 CCGAGCGAGCGCGCTGACAGCTCTTCCGAGCTCGGCTCCCGCAACATCTCTCTT 2443
DB 55 CCGAACCAACCGCTGACAACTTCCCAACTCGCGCTCCCGCAACATCTCTCTT 2

RESULT 6
ABO22411
ID ABO22411 standard; DNA; 1026 BP.
XX
```

AC AB022411;
 XX 12-JUL-2002 (first entry)
 DT
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 9002.
 DE
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP010074.
 XX
 PR 01-SEP-2000; 2000DE-01043826.
 PR 05-SEP-2000; 2000DE-01044543.
 XX
 PA (EPIC-) EPIDEMIOLOGICS AG.
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridization to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridized to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP/s); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. AB013410-
 CC AB054121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention
 CC
 SQ Sequence 1026 BP; 329 A; 429 C; 128 G; 140 T; 0 U; 0 Other;
 XX
 XX
 Query Match 19.0%; Score 604; DB 6; Length 1026;
 Best Local Similarity 80.5%; Pred. No. 9.1e-114;
 Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;
 XX
 QY 1490 CACGATCTACACGCGCTGTACATGAGCCGCTGACCCCTCGCCCAAGACTCGCGCG 1549
 DB 78 CAAATCTACACGCGCTGTATTAACGCTTAACCCCTCGCCCAAGACTCGCGCG 137
 QY 1550 CTGCGAAGTGTGCGCGCGCGCGCTACGCGCGCGCTCTTTCTGAGGACCAATT 1609
 DB 138 CTGCGAAGTGTGCGCGCGCGCGCTACGCGCGCGCTCTTTCTGAGGACCAATT 197
 QY 1610 CGTGAAGCACTACGCGCGCGCGCTTCCACCCGCGCGCGCGCGCGCGCGCG 1669
 DB 198 CGTGAAGCACTACGCGCGCGCGCTTCCACCCGCGCGCGCGCGCGCGCGCG 255

QY 1670 TACGGAACGAGCGTGTGCGCGCAACCAAGGCTGTGTGCGCGCGCGCGCGCGCG 1729
 DB 256 TACGGAACGAGCGTGTGCGCGCAACCAAGGCTGTGTGCGCGCGCGCGCGCGCG 315
 QY 1730 GGGGCGCGCGCTGTGCGCGCGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCG 1789
 DB 316 GAAAGCG 374
 QY 1790 GGGCTGTGCG 1849
 DB 375 GACCTGTGCG 434
 QY 1850 CG 1909
 DB 435 CG 493
 QY 1910 CGGCTACTACG 1969
 DB 494 CGACTACTACG 553
 QY 1970 CAAGTGTGCG 2029
 DB 554 CAAATGTGCG 612
 QY 2030 CGGCG 2089
 DB 613 CGAGCG 672
 QY 2090 GCG 2149
 DB 673 ACCG 731
 QY 2150 GACG 2209
 DB 732 AACG 791
 QY 2210 GCGGAGCG 2269
 DB 792 ACGAAGCG 851
 QY 2270 CGAGGCG 2329
 DB 852 CGAATGTGCG 911
 QY 2330 TGGCTGTGCG 2389
 DB 912 TAACTGTGCG 971
 QY 2390 CCCG 2443
 DB 972 CCCAAGCG 1025
 XX
 XX
 RESULT 7
 ACH73791
 ID ACH73791 standard; DNA; 515 BP.
 XX
 AC ACH73791;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon probe #6986.
 DE
 KW Human; probe; ss; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 OS Homo sapiens.
 PN US2003194704-A1.
 XX
 PD 16-OCT-2003.
 PF 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.
 PR
 PA (PENNY/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX
 PI Penn SG, Rank DR, Hanzel DK;
 DR WPI: 2004-119264/12.
 XX
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 PS
 PS Claim 15; SEQ ID NO 6986; 80bp; English.
 CC The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridises under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 SQ Sequence 515 BP; 87 A; 190 C; 106 G; 132 T; 0 U; 0 Other;
 Query Match 16.2%; Score 515; DB 12; Length 515;
 Best Local Similarity 100.0%; Pred. NO. 1.3e-95;
 Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 181 ||||| CCAGCTCGGCTCCGCCACACATCTCTTGCGCAACCACTATTATTGACCTCGATG 240
 QY 2476 GCGCTTGGACGAGTAATAGTGCAGGTCTCCGAGGCTGATTTTAACCTTTTGGACAGCA 2535
 DB 241 GCGCTTGGACGAGTAATAGTGCAGGTCTCCGAGGCTGATTTTAACCTTTTGGACAGCA 300
 QY 2536 GTCTGTGAATTACTGACCGACCTTCAACTTTCGTGAAACCTTTGGTTTCTTACTT 2595
 DB 301 GTCTGTGAATTACTGACCGACCTTCAACTTTCGTGAAACCTTTGGTTTCTTACTT 360
 QY 2596 ACTCTTCTTGTGAGATTATCTCTTCAATTCCTCCCTCGTCTTTCTTACTT 2655
 DB 361 ACTCTTCTTGTGAGATTATCTCTTCAATTCCTCCCTCGTCTTTCTTACTT 420
 QY 2656 CCTACTTCTTCTTTCTTGTGAATGAAACTCTTCACTTTAGAGACCTTGGCGATCTGTCA 2715
 DB 421 CCTACTTCTTCTTTCTTGTGAATGAAACTCTTCACTTTAGAGACCTTGGCGATCTGTCA 480
 QY 2716 GGACAGACGATTCGACCCGCAAGTCTGGGCT 2750
 DB 481 GGACAGACGATTCGACCCGCAAGTCTGGGCT 515
 RESULT 8
 ID ABT07220/C
 ID ABT07220 standard; DNA; 514 BP.
 AC ABT07220;
 DT 07-NOV-2002 (first entry)
 XX
 DE Human CpG-island rich DNA sequence SEQ ID No 37.
 XX
 KW CpG island; methylated; methylation-sensitive restriction enzyme; cancer;
 KW malignant cell; tumour; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200260318-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2002; 2002WO-US003077.
 XX
 PR 31-JAN-2001; 2001US-00775398.
 XX
 PA (OHIS) UNIV OHIO STATE RES FOUND.
 XX
 PI Plass C;
 XX
 XX WPI: 2002-627436/67.
 DR
 PT Identifying methylated CpG islands in malignant cells, useful for
 PT diagnosing or classifying cancer, comprises the use of infrequent
 PT cleaving, methylation-sensitive restriction enzymes and gel
 PT electrophoresis.
 XX
 PS Claim 17; Page 86; 133pp; English.
 XX
 CC The invention relates to a method for identifying CpG islands which are
 CC preferentially methylated in malignant cells, which comprises obtaining
 CC genomic DNA from malignant and non-malignant cells, digesting this DNA
 CC with infrequently-cutting, methylation-sensitive restriction enzymes to
 CC form restriction fragments, electrophoresing these fragments and
 CC comparing the intensity of these patterns to control cell restriction
 CC fragments. The method is useful in diagnosing cancer, in classifying
 CC tumours, in providing prognoses for cancer patients, and in identifying
 CC new DNA clones within a library based on the methylation status of CpG
 CC dinucleotides. The isolated polynucleotides and oligonucleotides can also
 CC be used as probes to identify genes whose expression is increased or
 CC decreased in cancerous tissues. This polynucleotide sequence represents a
 CC CpG-island rich DNA sequence relating to the invention


```
Db 413 GCGCTTAATTTTATTTGGCGAGAGGTCCAGGGTTTGTCGTGAGCGTTTCGTCTTG 354
Qy 2091 CCGCCCGCGCGCGCGAGAGCCCAAGCCCAAGACTGTCCGATTCCAGCTGATCGAG 2150
Db 353 TCGTTC-CCGTGCTCGAGAGCGTTAAGTTTGAATTTGTTAGTTTATTTAGTTGAG 295
Qy 2151 ACGCCCTCTCGATCAAGTCATCGACTCCAGCTCGGGGATTATAGAGAGGCCAAG 2210
Db 294 ACGTTTATTTTTCATTAAGTTTATCAATTTTACGATTGCGGGATTACAGTAGGTTAAG 235
Qy 2211 CGAGAGCGGATCTCCCGCGCCACACAGCCCGTGTCCGAGAGTTCCGTCCCGCTCAAGAGC 2270
Db 234 CGAGAGCGGATTTCTCGGTGATACGTTGTTGTTGTTGAGAGTTGTTTTCGTTTAAGGC 175
Qy 2271 GAGGTCTGCGCCAGCGGAGCTGCGAGAAAGAACTGCGCCAGCATTAAGCGCTACTAT 2330
Db 174 GAGGTGTTGTTTACCGGAGATTGCGAAGAAATTCGTTAAGATATTAGCGGTTATTAT 115
Qy 2331 GCGTCTACTCGACAGCTAGGCGCCCGCTCGCGCGCGCGCGCGCGCGCGAGACC 2390
Db 114 GGTTTTATCGTAAGTTAAGTTAGGTCGTTTGTTCGTTGCGTTTCGTCGAGTTT 55
Qy 2391 CCAGCCAGCCCTCAACAGCTCTTCCCAAGCTCCGCTCCCACTCCCTCTTG 2444
Db 54 TTAGTTAGTTTATTAAGTTTATTTTATTTTATTTTATTTTATTTTATTTT 1
```

```
RESULT 10
ABQ22412
ID ABQ22412 standard; DNA; 1026 BP.
XX
AC ABQ22412;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 9003.
XX
KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KM drug; side effect; cancer; central nervous system; cardiovascular;
KM gastrointestinal; respiratory system; single nucleotide polymorphism;
KM SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
FN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PE 01-SEP-2001; 2001WO-BP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
```

degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. ABQ13410-CC ABQ54121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention

Seq Sequence 1026 BP; 146 A; 128 C; 311 G; 441 T; 0 U; 0 Other;

Query Match 14.7%; Score 469.6; DB 6; Length 1026;
Best local Similarity 71.7%; Pred. No. 3.2e-86;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;

```
Qy 1491 ACGATCTACACCGGCTGTGACATGACCCGCTGACCCCTGCCCAAGACTCGCGCGC 1550
Db 79 AGGATTATATATCGGTTGTGATATGATCGTTTGAATTTTCTGTTACGATTCTGCGGT 138
Qy 1551 TCGCAGATCGTCCCGCGCGCGCTACGCGCATGCGCCGCTCTTCTGACAGACAGTTTC 1610
Db 139 TCGTAGATCGTGTTCGCGGTTTCGTTACGTTATGATCGGTTTCTTGTAGAGATTAGTTC 198
Qy 1611 GTGAGCAACTAGCCCAAGCCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGGT 1670
Db 199 GTGAGTATATTAAGTTAAGTTAGTTGTTTATTCGCGCGCGCG--CGCGGTTTCGCGTCCGCT 256
Qy 1671 ACGAGACGCGACGCGCGCGCACACCAACGCGGCTGCTGTGCGCGCGAGAGCCGAGACCCG 1730
Db 257 ACGAGTCTAGCGGTGCTGATATTAACGCGTTGTTGCTGTGATAGTGAAGATTTCG 316
Qy 1731 GCGCGCGCTTCCGCGCAACGCTGTTTGTGACGCGCGCGCAACAACGCGTGAAGTTTCGCG 1790
Db 317 GCGCGCG--TTTCGTGTGTAACGTTGTTGTGACGTCGTTAATATCCGTTGAGTTTGGCG 375
Qy 1791 GCTTCGCGCTATGACAGCGCCACGACTTCGCGGCGCAACGCGCGCACGCTGCTCTTAC 1850
Db 376 GTTTCGTTTATGATACGCTTACGGAATTCGCGGATTAACGCGGATTAAGTTTATTTTAC 435
Qy 1851 GCGCGCGCGCGCGGAGAGCGCTGCGCGAGGCTGAGGCTGACGTCGCGCGCGCTC 1910
Db 436 GCGCGCGCGCGCGTGAAGCGTGT-CGTTGAGGTTGATGATGATGATGATGATGATGATG 494
Qy 1911 GCGTACTACGCGCAACCGCTCGGCGTGGGCGCGCGCAAGTCCCGCAAGTACTGCGGCAAC 1970
Db 495 GGTATTATACGTCGATTCGTTGCGGTTGGGCGCTTGTGATTTTGTGATGATGCGGATTT 554
Qy 1971 AAGTCGCGCTCGTGTGCTGCTGCTGCGCAACAGCGCGCGCGCGCGCGCATGCGC 2030
Db 555 AAGTCGCGGTTGCGTGTG-TTTTGTGTTGTTTATAGCGTCGCGCGCGCATGCGTTC 613
Qy 2031 GCGCGCAATCCCTACTCGGCGGAGAGAGCGAGGCGCTGCGCGCGAGCGCTGCGCGCTG 2090
Db 614 GCGCTTAATTTTATTTTATTTGGCGAGAGGTCAGAGGTTTGTGTCGAGCGTTCTGTTG 673
Qy 2091 CCGCCCGCGCGCGCGCGAGAGCGCAAGCCCAAGACTGTTCGATTCAGCTGATCGAG 2150
Db 674 TCGTTC-CCGTGCTCGAGAGCGTTAAGTTTGAATTTGTTTGAATTTAGTTGATGAG 732
Qy 2151 ACGCCCTCTCGATCAAGTCATCGACTCCAGCACTCGGGGATTATAGAGAGGCCAAG 2210
Db 733 ACGTTTATTTTTCATTAAGTTTATGATTTTATGCAATTCGCGGATTTACGATAGTTAAG 792
Qy 2211 CGAGAGCGGATCTGCGCGCGCGAGACCGCGCTGTCCAGAGTTGTCGCCGCTCAAGAGC 2270
Db 793 CGAGAGCGGATTTGTCGCTGATACGTTCTGTTCGAGATTCGTTTCTTTTAAGAGC 852
Qy 2271 GAGGTGCTGCGCCAGCGGAGCTGCGAAGAAATGCGCAAGACATTAGCGGCTACTAT 2330
```


mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors

SQ Sequence 388 BP; 99 A; 90 C; 96 G; 102 T; 0 U; 1 Other;

Query Match	12.1%	Score 384.6	DB 3	Length 388
Best Local Similarity	99.7%	Pred. No. 6.5e-69		
Matches 384; Conservative	1	Mismatches 0	Indels 0	Gaps 0

Oy	190	TTTGGCTGGTTGAAGGCTTCTGTCTAGTGAAGGGGGCTGTGGATTCTTCAATTATGATA	249
Db	4	WTGCGGTGAAGTGCTTCTGTCTAGTGAAGGGGGCTGTGGATTCTTCAATTATGATA	63
Oy	250	AATAGCACTTTAAAAACGAGGACGGGAGGGCAGTGTTCAGGTTCTAGAGCTATGCAGC	309
Db	64	AATAGCACTTTAAAAACGAGGACGGGAGGGCGAAGTTCAGGTTCTAGAGCTATGCAGC	123
Oy	310	TGAGCACTGCTTTCTCCTTCTATCATGCTCTCCAAAGAAATTTCTCAATGTGAGACGA	369
Db	124	TGAGCACTGCTTTCTCCTTCTATCATGCTCTCCAAAGAAATTTCTCAATGTGAGACGA	183
Oy	370	GCTACCCACATTCGAGGGGATCCGAGCTTGTCTTGACAGATCATGCCATTATCTGACCA	429
Db	184	GCTACCCCAATTCGAGGGGATCCGAGCTTGTCTTGACAGATCATGCCATTATCTGACCA	243
Oy	430	CTGACAACTGTGAGAGAAATTCACTTTGAAAAAATTAACGAGGGGATGACCAATCAGT	489
Db	244	CTGACAACTGTGAGAGAAATTCACTTTGAAAAAATTAACGAGGGGATGACCAATCAGT	303
Oy	490	CAGATACAGACAATTTCTCTGACTCCAGAGACTCACAGGGGACGTCCAGAGAAAGTAAAC	549
Db	304	CAGATACAGACAATTTCTCTGACTCCAGAGACTCACAGGGGACGTCCAGAGAAAGTAAAC	363
Oy	550	TCTCTCTCTGTCTTTGACGGGGGTCTC	574
Db	364	TCTCTCTCTGTCTTTGACGGGGGTCTC	388

RESULT 13

AAZ46477
ID AAZ46477 standard; cDNA; 2494 BP.

AC AAZ46477;

DT 13-MAR-2000 (first entry)

Human T-box transcription factor protein, MTbx encoding cDNA.

KW T-box transcription factor protein; MTbx; immune system disease; HIV
 KW leukemia; inflammatory disease; asthma; rheumatoid arthritis; human;
 KW developmental disorder; cardiovascular; transcriptional regulation;
 KW myocardial ischemia; gene mapping; chromosome 3p23-p24; ss

OS Homo sapiens.

Key	Location/Qualifiers
FH	

FT	CDS	164. .1717
----	-----	------------

```
FT /product= "Mtbx"
```

PN WO9966025-A2.

PD 23-DEC-1999.

PF 15-JUN-1999; 99WO-US013457.

PR 16-JUN-1998; 98US-0089467P.

PR 09-NOV-1998; 98US-00188811.

XX

XX

XX

DR P-PSDB; MAY49

PT Nucleotides encoding

XX

XX XX

This cDNA encodes a human T-box transcription factor protein, MTbx. The polypeptide can be administered therapeutically, especially by expressing encoding polynucleotides, to treat disorders characterized by aberrant polypeptide expression or activity, e.g. immune system diseases (e.g. HIV, leukemia, etc.), chronic inflammatory diseases (e.g. asthma, rheumatoid arthritis etc.), developmental disorders, cardiovascular disorders, or other disorders arising from improper transcriptional regulation. They are especially useful to treat conditions characterized by loss of tissue integrity relating to disease and/or injury e.g. HIV, hibernating myocardium during myocardial ischemia etc. It can be used to diagnose diseases related to polypeptide expression or activity. The polynucleotide may also be used to produce non-human transgenic animals and to produce probes useful to detect related sequences, or for gene mapping. The Mtbx gene was mapped to chromosome 3p23-p24

Sequence 2494 BP; 651 A; 639 C; 600 G; 594 T; 0 U; 10 Other;

Query Match	11.8%	Score 374.6;	DB 3;	Length 2494;
Best Local Similarity	68.8%	Pred. No. 1.2e+66;		
Matches 515; Conservative	0;	Mismatches 234;	Indels 0;	Gaps 0;

QY	912	GCACAGGCTGACCTGTGACACAGGCCCCCTTTGGCTGAAATTTACCCGGCACCAACCGAG	971
Db	398	GCCACGCTACCTGTGACACCGGCTCTGTGGCTCAATATTCACCGCACCAACTGAG	457
QY	972	ATGATCATCACCAACACAGGAGGGCGCATGTTCTCTTTTAACTTTAACTTTCTGGT	1031h
Db	458	ATGATCATTTAAGAAACAGGCGAGGTCTTCTTTCTTGAGCTTCATATAACGGA	517
QY	1032	CTCGATCCCAAGGCTCATTAACAATTTTGTGATGTGATTTTGGCGGATCCCAATAC	1091h
Db	518	CTCATCCCACTGCCCCACTACAAATGTGTCTGTAAGGGTGTGTGGCGGACCCCAACAC	577
QY	1092	TGAGAGTTTCAAGGAGGCAAAATGGGTTCTTTGCGGCAAAAGCGGACACAATGTGTCAAGA	1151h
Db	578	TGGCGCTTCCAGGGGGGCAAAATGGGTACTGTGTGCAAAAGCGGACAAATATCATGACGGC	637
QY	1152	AATCGGCTTAATGTCATCCGAGATTTCCCCAACACTGGGGTCTCATCTGANTGCGCCAAAGA	1211h
Db	638	AACAAAATGTATGTTCACCCAGAGCTCTTAATCTGGTTCCCACTGGATGAGACAGAG	697
QY	1212	ATCTCTTTTGGAAATTTAAACTTACGAAACAAGAGGTTCAAAATTAACAATGGGAG	1271h
Db	698	ATTTCTATTGGGAAATTTAAACTCACCAATTAACAAAGGCGCAAAATTAACAACAACCCAG	757
QY	1272	ATGTGTGTTTACAGTCTCTTGACAGATACAGGCCCGCTGTGATGTGTGGAAGTGAAC	1331h
Db	758	ATGATAGTCTACATCTCTTACAAATTAACAACCAACCCGAGCTGATATTTGTTGAAGTTACA	817
QY	1332	GAGACGGCAACGAGAGACCTAGCCAGCTCGGCGCGCGTGTGACAGCTTCACTTTCCCTGAG	1391h
Db	818	GAGGATGCGCTGGAGACTTAATAGAGCTCTTCAAAAGACCCGACCTTTTAACTTCTTCAGAA	877

```

QY      1392 ACTCAGTTCATGCGCGTACCGCTACCAAGAACAGGATATTACACAACTGAAATAGAT 1451
DB      878 ACGCAATTCATTGACAGTGAAGTCTCTACCAAAAACCGGATATTACCACTAAAGATTGAT 937
QY      1452 CACAAACCTTTTGGAAAAAGATTGGGATATTATGTGACAGATCTACACGGGCTGTAC 1511
DB      938 CATACACCTTTTGGAAAAAGCTTCAGAGCAACTATGATTCATGTACACCGCTTCAGAA 997
QY      1512 ATGACCGCGCTACCGCCCTCGCCCAAGCACTCGCGCTCGCAGATGTCGCCGGGGCC 1571
DB      998 AATGACAGGTAACTTCATCTCTCCACGATTTCTCTAATGCCATCAGATTGTCCCTGGA 1057
QY      1572 CGCTACGCCATGCGCGGCTCTTTCTCTGACAGACAGTTCTGAGCAACTACGCCAAGGCC 1631
DB      1058 GGTGCGTACGGCGTTCAATCTCTTCCCGGAGCCCTTTGTCAACACTTACTCAAGGCC 1117
QY      1632 CGCTTCACCGCGCGCGCGCGCGCGCGGCC 1660
DB      1118 CGCTATTATATGCGAGAGAACCGTGC 1146

RESULT 14
ADP28768
ID      ADP28768 standard, DNA; 3465 BP.
XX
AC      ADP28768;
XX
DT      12-AUG-2004 (first entry)
XX
DE      Human secreted protein encoding sequence SEQ ID #766.
XX
KW      Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX      cancer; inflammatory; immune; ds; human secreted protein.
OS      Homo sapiens.
XX
PN      WO2004035732-A2.
XX
PD      29-APR-2004.
XX
PF      28-AUG-2003; 2003WO-US026780.
XX
PR      29-AUG-2002; 2002US-0406576P.
PR      29-AUG-2002; 2002US-0406579P.
PR      29-AUG-2002; 2002US-0406585P.
PR      29-AUG-2002; 2002US-0406588P.
PR      29-AUG-2002; 2002US-0406608P.
PR      29-AUG-2002; 2002US-0406611P.
PR      29-AUG-2002; 2002US-0406612P.
PR      29-AUG-2002; 2002US-0406616P.
PR      29-AUG-2002; 2002US-0406640P.
PR      29-AUG-2002; 2002US-0406642P.
PR      29-AUG-2002; 2002US-0406646P.
PR      29-AUG-2002; 2002US-0406653P.
PR      29-AUG-2002; 2002US-0406655P.
PR      29-AUG-2002; 2002US-0406666P.
PR      17-SEP-2002; 2002US-0410946P.
PR      17-SEP-2002; 2002US-0410947P.
PR      17-SEP-2002; 2002US-0410948P.
PR      17-SEP-2002; 2002US-0410949P.
PR      17-SEP-2002; 2002US-0410953P.
PR      17-SEP-2002; 2002US-0410957P.
PR      17-SEP-2002; 2002US-0410958P.
PR      17-SEP-2002; 2002US-0410959P.
PR      17-SEP-2002; 2002US-0410960P.
PR      17-SEP-2002; 2002US-0410961P.
PR      17-SEP-2002; 2002US-0410962P.
PR      17-SEP-2002; 2002US-0411019P.
PR      17-SEP-2002; 2002US-0411022P.
PR      17-SEP-2002; 2002US-0411023P.
PR      17-SEP-2002; 2002US-0411024P.
PR      17-SEP-2002; 2002US-0411032P.

```

```

PR      17-SEP-2002; 2002US-0411035P.
PR      17-SEP-2002; 2002US-0411037P.
PR      17-SEP-2002; 2002US-0411041P.
PR      17-SEP-2002; 2002US-0411045P.
PR      17-SEP-2002; 2002US-0411046P.
PR      17-SEP-2002; 2002US-0411048P.
PR      17-SEP-2002; 2002US-0411052P.
PR      17-SEP-2002; 2002US-0411055P.
PR      17-SEP-2002; 2002US-0411073P.
PR      17-SEP-2002; 2002US-0411082P.
PR      17-SEP-2002; 2002US-0411101P.
PR      17-SEP-2002; 2002US-0411111P.
PR      18-APR-2003; 2003US-0463700P.
PR      18-APR-2003; 2003US-0463708P.
PR      18-APR-2003; 2003US-0463716P.
PR      18-APR-2003; 2003US-0463732P.
PR      02-MAY-2003; 2003US-0467199P.
PR      02-MAY-2003; 2003US-0467201P.
PR      02-MAY-2003; 2003US-0467203P.
PR      02-MAY-2003; 2003US-0467230P.
PR      19-MAY-2003; 2003US-0471306P.
PR      19-MAY-2003; 2003US-0471336P.
PR      22-MAY-2003; 2003US-0472420P.
PR      22-MAY-2003; 2003US-0472430P.
PR      09-JUN-2003; 2003US-0476609P.
PR      09-JUN-2003; 2003US-0476641P.
PR      08-JUL-2003; 2003US-0485218P.
PR      08-JUL-2003; 2003US-0485223P.
PR      08-JUL-2003; 2003US-0485224P.
PR      14-JUL-2003; 2003US-0486325P.
PR      14-JUL-2003; 2003US-0486446P.
PR      14-JUL-2003; 2003US-0486480P.
PR      15-JUL-2003; 2003US-0486891P.
PR      15-JUL-2003; 2003US-0486960P.
PR      08-AUG-2003; 2003US-0493341P.
PR      08-AUG-2003; 2003US-0493370P.
PR      08-AUG-2003; 2003US-0493573P.
PR      08-AUG-2003; 2003US-0493577P.
XX
XX      (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PA      Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI      Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linemann T;
PI      Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX      WPI; 2004-348438/32.
XX
PT      New nucleic acid molecule for diagnosing, preventing or treating diseases
PT      such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX      genetic, bacterial and viral diseases.
XX
XX      Claim 1; SEQ ID NO 766; 428bp; English.
XX
XX      The present invention relates to an isolated nucleic acid molecule
XX      encoding a polypeptide which is believed to be cytostatic,
XX      antinflammatory, immunosuppressive, antibacterial and virucidal. The
XX      composition and methods are useful for diagnosing, preventing and
XX      treating diseases such as proliferative (e.g. cancer), inflammatory,
XX      immune, metabolic, genetic, bacterial and viral diseases. The present
XX      sequence represents a human secreted protein encoding sequence. The
XX      present sequence is available on WIPWEB and is not in the specification.
XX
SQ      Sequence 3465 BP, 847 A, 1005 C, 912 G, 701 T, 0 U, 0 Other;
XX
Query Match      11.8%; Score 374.6; DB 12; Length 3465;
Best Local Similarity 68.8%; Pred. No. 1.3e-66;
Matches 515; Conservative 0; Mismatches 234; Indels 0; Gaps 0;
QY      912 GCACAGGTATCTGTGCAACAGCGCCCTTTGGCTGAATTTACCGGACCAAGGAG 971
DB      952 GCCACAGTCTACTGTGCAACAGCGGCTCTGTGGCTCAATTCACCGGACCAACTGAG 1011
QY      972 ATGATCATCACCAACAGGAGGCGCATGTTCTTTTAAGTTTAACTTTCTGGT 1031

```

```

Db      1012 ATGATCATTACGAAACAGGCGAGCGCAATGTTCTCTTCTTGAGCTTCAACATTAACGGA 1071
Qy      1032 CTGATCCCAAGCGCTCATTAACAATTTTGTGATGTGATTTGGCGGATCCCAATAC 1091
Db      1072 CTCATCCCACTGACCACATACATATGTGTGTAGAGGTGTCTGGCGGACCCCAACAC 1131
Qy      1092 TGGAGGTTTCAAGAGGCAATGGGTCTTGGCGCAAGCGGACCAATGTGCAAGA 1151
Db      1132 TGGCCCTTCAGAGGCGCAATGGGTGACCTGTGGCAAGCGGACCAATGACGAGGC 1191
Qy      1152 AATCGGCTTATATGATCCGATCCGATTCGCCCAACACTGGGGCTCACTGTATGCCAAGA 1211
Db      1192 AAGCAAAATGTATGTTCAACCAAGCTCTCTAATACGTGTTCCACGTGATGACAGGAG 1251
Qy      1212 ATCTCTTTTGAATAATTAACCTTACGAACAACAAGAGCTTCAATATACATGGGAG 1271
Db      1252 ATTTATTCGGGAAATTAACCTCACACATATACAAAGGCGCAATATACACACACGAG 1311
Qy      1272 ATGGTGGTTTACAGTCTTGCAAGATACCAAGTCCGCGCGCTGCATGTGGTGAAGTGAAC 1331
Db      1312 ATGATATGTTTACAAATCTTATACAAATACCAACCCGACTGATATGTTGAAGTTACA 1371
Qy      1332 GAGGACGCGACGAGGACACATAGCCAGCCGCGGTGACAGCGTTCACTTCCCTGAG 1391
Db      1372 GAGGATGGCGTGGAGGACTTGAATAGCCCTCAAAAGACCCAGACTTTTACCTTCAAGA 1431
Qy      1392 ACTCATTTATCGCGCTACCGCTTACCAAGAACAGGATATTAACAACATAAGTATAGT 1451
Db      1432 ACGCAATTCATGTGACGTGACTGCTTACCAAAACCGGATATTACTCACTAAAGATTGAT 1491
Qy      1452 CACAACCCCTTTGCAAAAGAGATTGGGGATTAATTAAGACAGATCTACACCGGCTGTGAC 1511
Db      1492 CATTAACCCCTTTGCAAAAGGCTTCAAGACAACTATGATTCATGTACACCGCTTCAAGA 1551
Qy      1512 ATGACCGCGCTACCCCTGCGCAACGACTGCGCGCTGCAAGATCGTGC CGGAGCC 1571
Db      1552 AATGACAGGTAACTCATCTCCACAGGATTTCTCCAGATCCCATCAAGTTGCTCCGGA 1611
Qy      1572 CGCTACGCGCATGCGCGCTTTTCTTGACAGACAGTGTGTAGACAATACCCAAAGCC 1631
Db      1612 GGTGGGTATGCGGTTCATCTCTTCCGAGGCCCTTGTCAACACTTACCTCAAGGC 1671
Qy      1632 CGCTTCACACCGGCGCGGCGGCGGAGCC 1660
Db      1672 CGCTATTAATATGCGAGAAACCGTGCC 1700

```

RESULT 15

AA039695
ID AA039695 strand: DNA; 389 BP.

```

AC      AA039695;
XX      25-MAR-2003 (revised)
DT      20-MAY-1993 (first entry)
XX      DE      Expressed Sequence Tag human gene marker EST00360.
XX      KW      expressed sequence tag; human genome project; chromosome;
XX      KW      human gene sequencing; PCR mapping; somatic cell hybrids;
XX      KW      sublocalisation; gene tagging; tissue typing.
OS      Synthetic.
XX      PN      WO9300353-A1.
XX      PD      07-JAN-1993.
XX      PF      19-JUN-1992; 92WO-US005222.
XX      PR      20-JUN-1991; 91US-00716831.
XX      PR      12-FEB-1992; 92US-00837195.

```

```

XX      PA      (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX      PI      Venter JC, Adams MD;
XX      DR      WPI; 1993-036325/04.
XX      PT      Particular expressed sequence tags from human CDNA - corresponds to
XX      PT      transcriptions of genes, useful for tagging genes, mapping
XX      PT      chromosomes and tissue typing.
XX      PS      Claim 3; Page 105; 199pp; English.
CC      CC      This sequence represents an EST (expressed sequence tag) ESTs are markers
CC      CC      for human genes actually transcribed in vivo. Unlike the random genomic
CC      CC      DNA sequence tagged sites (STSs), ESTs point directly to expressed genes.
CC      CC      The use of ESTs could facilitate the tagging of most expressed human
CC      CC      genes within a few years at a fraction of the cost of complete genomic
CC      CC      sequencing. Using PCR primers AA039419-039580 (sequences designed from
CC      CC      the ESTs) sublocalisation of an EST can be achieved with panels of
CC      CC      fragments from specific chromosomes or pools of large genomic clones in
CC      CC      an analogous manner. This sequence represents EST00360. (Updated on 25-
CC      CC      MAR-2003 to correct PN field.)
SQ      SQ      Sequence 389 BP; 97 A; 81 C; 71 G; 139 T; 0 U; 1 Other;
Query Match      11.4%; Score 363; DB 2; Length 389;
Best Local Similarity 99.2%; Pred. No. 1.7e-64;
Matches 385; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
Qy      2775 TCTAGAACCTTGACCCACCCAGCGCGTCTTCTTATCCCGAGTGAATGATGATGGA 2834
Db      1 TCTAGAACCTTGACCCACCCAGCGCGTCTTCTTATCCCGAGTGAATGATGATGGA 60
Qy      2835 TGAATGTAAGGATTAATTAATTTTGTGGAACAAGCTGTGAATGATTTGACATAG 2894
Db      61 TGAATGTAAGGATTAATTAATTTTGTGGAACAAGCTGTGAATGATTTGACATAG 120
Qy      2895 TGTTAATTAATTAATTAATTAATTTTGTGGAACAAGCTGTGAATGATTTGACATAG 2954
Db      121 TGTTAATTAATTAATTAATTAATTTTGTGGAACAAGCTGTGAATGATTTGACATAG 180
Qy      2955 TGTTAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 3014
Db      181 TGTTAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 240
Qy      3015 TTTCTCTCTTTTAATTTCTTGTGAGATTAATTTCTTAAGAGGCTTAAGAAACATGAAT 3074
Db      241 TTTCTCTCTTTTAATTTCTTGTGAGATTAATTTCTTAAGAGGCTTAAGAAACATGAAT 300
Qy      3075 ACTGAGTAGT-GATGGGTTTCCCACTTCTCTCAATCGGTGATGAATAATTAAT- 3132
Db      301 ACTGAGTAGTGAATGGGTTTCCCACTTCTCTCAATCGGTGATGAATAATTAAT- 360
Qy      3133 GTGCCCTAATGCACAAATAGCTAAG 3160
Db      361 GTGCCCTAATGCACAAATAGCTAAG 388

```

Search completed: December 20, 2005, 08:14:21
Job time : 1811 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 07:07:03 ; Search time 12029 Seconds
(without alignments)
12392.023 Million cell updates/sec

Title: US-10-649-400-3

Perfect score: 3186
Sequence: 1 caggtgattatcctaataa.....cacccaacaccttaag 3186

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1002.6	31.5	1116	5	BX418703	BX418703 BX418703
2	935.4	29.4	957	5	BX441549	BX441549 BX441549
3	897	28.2	921	5	BX420367	BX420367 BX420367
4	823	25.8	930	3	B1912693	B1912693 603176344
5	811.4	25.5	824	5	BX418656	BX418656 BX418656
6	809	25.4	820	3	B1915063	B1915063 603177258
7	801.6	25.2	872	3	B1914330	B1914330 603182311
8	798.2	25.1	1169	3	B0067178	B0067178 AGENCOURT
9	789.8	24.8	1044	3	BM922065	BM922065 AGENCOURT
10	766.6	24.1	945	3	B1914132	B1914132 603182077
11	761	23.9	966	3	B1754056	B1754056 603027611
12	760	23.9	831	1	AM003023	AM003023 wq61d04.x
13	749.6	23.5	910	7	CN643299	CN643299 ILLUMIGEN
14	722.4	22.7	768	8	DR001136	DR001136 TC115971
15	720	22.6	910	7	CN805179	CN805179 ILLUMIGEN
16	716	22.5	748	3	B1917426	B1917426 603184973
17	710	22.3	991	3	B1757271	B1757271 603030777
18	699.8	22.0	791	3	B1918360	B1918360 603182871
19	695.6	21.8	779	3	B1820633	B1820633 603034884
20	692.8	21.8	740	3	B1917781	B1917781 603183881
21	692.8	21.7	777	6	CD355525	CD355525 UI-M-FVO-
22	691.4	21.7	1154	3	BM807553	BM807553 AGENCOURT

23	687.6	21.6	772	3	B1823704	B1823704 603040821
24	682.4	21.4	739	3	B1918597	B1918597 603176527
25	681.8	21.4	1028	3	BM548197	BM548197 AGENCOURT
26	664.8	20.9	712	6	CA326858	CA326858 UI-M-FVO-
27	664.6	20.9	749	5	BQ443529	BQ443529 UI-M-ENO-
28	658.4	20.7	787	3	B1917926	B1917926 603181686
29	642.2	20.2	703	6	CA326808	CA326808 UI-M-FVO-
30	640.8	20.1	716	6	CA15364	CA15364 UI-M-FVO-
31	637.6	20.0	716	3	B0179248	B0179248 UI-M-ENO-
32	625	19.6	674	6	CB244022	CB244022 UI-M-FVO-
33	616	19.3	639	3	B1819019	B1819019 603033141
34	613	19.2	765	3	B1912695	B1912695 603176346
35	610.6	19.2	1015	3	BM921945	BM921945 AGENCOURT
36	593.8	18.6	606	1	AM590019	AM590019 hg29e09.x
37	592.2	18.6	700	6	CF531585	CF531585 UI-M-FVO-
38	585.4	18.4	645	6	CB527374	CB527374 UI-M-FVO-
39	582.2	18.3	801	8	CB821360	CB821360 JGI_CAAK2
40	581	18.2	583	3	BP311133	BP311133 BP311133
41	560.4	17.6	582	3	BP228857	BP228857 BP228857
42	554.8	17.4	855	8	CB829223	CB829223 JGI_CAAK6
43	515.8	16.2	534	5	BX114402	BX114402 BX114402
44	508.8	16.0	541	1	A1201654	A1201654 gi41f08.x
45	506.8	15.9	566	7	CV562445	CV562445 UI-M-FCO-

ALIGNMENTS

RESULT 1
BX418703
LOCUS
DEFINITION
BX418703 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone
CS0DF008YF20 5-PRIME, mRNA sequence.
ACCESSION
BX418703
VERSION
BX418703.2 GI:46957805
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 1116)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On May 15, 2003 this sequence version replaced gi:30769503.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DF008YF20&cl=1851.f.
Location/Qualifiers
1..1116
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF008YF20"
/library="FETAL BRAIN"
/dev_stage="fetal"
/color_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

FEATURES

ORIGIN

Query Match	31.5%;	Score 1002.6;	DB 5;	length 1116;
Best Local Similarity	96.5%;	Pred. No. 1.9e-21;		
Matches 1062;	Conservative 14;	Mismatches 18;	Indels 7;	Gaps 5

QY	266	CAAGGAAAGGAGGGGCGAGTGTTCAGGTTGCTTAAGAGCTATGAGAGCTGAGAGCACTGGCTTTC	325
Db	2	CCGGAGTGGGAGGGGCGAGTGTTCAGGTTGCTTAAGAGCTATGAGAGCTGAGAGCACTGGCTTTC	61
QY	326	TCCTTTCAATCAATGCTCTCCCAAGAAATTTCTCAATGTGAGCAGCAGCTAACCCAAATTCAGG	385
Db	62	TCCTTTCAATCAATGCTCTCCCAAGAAATTTCTCAATGTGAGCAGCAGCTAACCCAAATTCAGG	121
QY	386	CGGATCCGAGCTTGTGCTTGCAAGATCATCCCAATTATCTGCACACTGACCAACTCTGGAGAG	445
Db	122	CGGATCCGAGCTTGTGCTTGCAAGATCATCCCAATTATCTGCACACTGACCAACTCTGGAGAG	181
QY	446	AAGTTCACTTTGAAAAAATTTACAGGGGGGATGACGAATCAGTCAGATTAAGACAAATTT	505
Db	182	AAGTTCACTTTGAAAAAATTTACAGGGGGGATGACGAATCAGTCAGATTAAGACAAACTGGAGAG	241
QY	506	TCCTGACTCCAAAGACTACCAAGGGGAGCTCCAGAGAACTCTCTCTGCTTGGGA	565
Db	242	TCCTGACTCCAAAGAGACTACCAAGGGGAGCTCCAGAGAACTCTCTCTGCTTGGGA	301
QY	566	CGGGGTCTCTGAGCTTGGTCAAGTTTGAGTGGCTCTGGTGAATGGCTTACCTCTCTC	625
Db	302	CGGGGTCTCTGAGCTTGGTCAAGTTTGAGTGGCTCTGGTGAATGGCTTACCTCTCTC	361
QY	626	TCAGTCCAGCAGCAGCAGCTGCGGCGCACTGCTCCAGTGCATGTTCCCGTACCCCGG	685
Db	362	TCAGTCCAGCAGCAGCAGCTGCGGCGCACTGCTCCAGTGCATGTTCCCGTACCCCGG	421
QY	686	CCAGCAGCGAATCCGGCGCACCCCGCTTCTCAATGGGAGCCTTACGCGCTTACATGGCCCA	745
Db	422	CCAGCAGCGAATCCGGCGCACCCCGCTTCTCAATGGGAGCCTTACGCGCTTACATGGCCCA	481
QY	746	CCACCCGGTCAATCAACAAGAGCCTTCAACAAGCCTCGTGTCCAACTCTGCGCGCAGG	805
Db	482	CCACCCGGTCAATCAACAAGAGCCTTCAACAAGCCTCGTGTCCAACTCTGCGCGCA-GG	540
QY	806	ATATCCCAACGGGCGGCTACCCCTTACCCACAGCAGTACGGGCACTCTTACCAAGAGACTCC	865
Db	541	ATATCCCAACGGGCGGCTACCCCTTACCCACAGCAGTACGGGCACTCTTACCAAGAGACTCC	600
QY	866	GTTCTTACCAAGTTCCTCTCCATCCAGCCGAGGCTGGTGCCTCGGCAAAAGCAAGGTATCCT	925
Db	601	GTTCTTACCAAGTTCCTCTCCATCCAGCCGAGGCTGGTGCCTCGGCAAAAGCAAGGTATCCT	660
QY	926	GTCGCAACAGGCCCCCTTGGCTGGAATTTTCAACGGGCAACAAAACGGAGATGATCATCAACAA	985
Db	661	GTCGCAACAGGCCCCCTTGGCTGGAATTTTCAACGGGCAACAAAACGGAGATGATCATCAACAA	720
QY	986	AACAGGGAAGGCGCATGTTTCTTTTAAATTGTTTAACTTTCTGGTCTCGATCCACGAGC	1045
Db	721	AACAGGGAAGGCGCATGTTTCTTTTAAATTGTTTAACTTTCTGGTCTCGATCCACGAGC	780
QY	1046	TCATTACAAATATTTTGTGTGAGATGTATTTTGGCGGATCCCAATCACTGAGAGTTTCAAGG	1105
Db	781	TCATTACAAATATTTTGTGTGAGATGTATTTTGGCGGATCCCAATCACTGAGAGTTTCAAGG	840
QY	1106	AGGCAAAATGGGTTTCCCTTGCGGCAAAAGCGGACACCAATGTGCACAGGAAATCGGGTCTATAT	1165
Db	841	AGGCAAAATGGGTTTCCCTTGCGGCAAAAGCGGACACCAATGTGCACAGGAAATCGGGTCTATAT	900
QY	1166	GCATCCGGAATTCGCCCAACACTGGGGGCTCACTGAGTGGGCCAAGAAATCTC-TTTTGGAA	1224
Db	901	GCATCCGGAATTCGCCCAACACTGGGGGCTCACTGAGTGGGCCAAGAAATCTC-TTTTGGAA	960
QY	1225	AATTAAACTTACAGAAACAACAAGAGCTTCAAAATTAACAATGGGAGAT-GGTGGTTTTA	1283
Db	961	AATTAAACTTAAAGAACAAACAAGAGAGCTTCAAAATTAACAATGGGAGATGGGTGGTTTTT	1020

QY	1284	CAGGCTCTTGGCAACACTGACGAGCCCCGGCTGATGTGTGGGAAGGACGAGAGCGACG	1343
Db	1021	CMGTCTTGTCCCAAGTTCMGCCCC	CGTGCATGTGTGAARTGA

QY	1344	GAGGACACTGACCGACGCCCGGC	1364
Db	1077	GRGGGAAMTATACMAAGCCCGGC	1097

RESULT 2	
BX441549	
LOCUS	BX441549 957 bp mRNA linear EST 05-MAY-2004
DEFINITION	BX441549 Homo sapiens FETL BPA1N Homo sapiens cDNA clone
	CSODF018YU12 5-PRIME, mRNA sequence.

VERSION	BX441549.2	GI:47038398
KEYWORDS	EST,	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE 1 (bases 1 to 957)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31018481

Genoscope - Centre National de Séquençage
2 rue Gaston Crémieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: segreg@genoscope.cns.fr, Web : www.genoscope.cns.fr
The Not I site was primed with a NotI-oligo (bp) primer. Five primed
end enriched, 1 double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cna/s=CS0DF018DE06QPlc=1851.f>.

```

FEATURES
    source
        location/Qualifiers
            1..957
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CSODF01BYJ12"
                /tissue_type="FETAL BRAIN"
                /dev_stage="fetal"
                /clone_lib="Homo sapiens FETAL BRAIN"
                /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dt) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

Query Match	Similarity	29.4%	Score 935.4	DB: 5	Length 957
Best Local	Similarity	99.6%	Pred. No. 3,66-215		
Matches	947	Conservative	1	Mismatches	2
				Indels	1
				Gaps	1
QY	387	GGATCCGAGCTTGTCTTGACGATCATCCATTTATCTTGACCACTGACCACTTGAGAGA	446		
Db	8	GGATCCGAGCTTGTCTTGACGATCATCCATTTATCTTGACCACTGACCACTTGAGAGA	67		
QY	447	AGTTCACCTTTGAAAAAATTAACGAGGGGGATGACGAATCAGTACGATACAGACAATTTT	506		
Db	68	AGTTCACCTTTGAAAAAATTAACGAGGGGGATGACGAATCAGTACGATACAGACAATTTT	127		
QY	507	CCTGACTCCAAAGGACTCACCGAGGGGAGCTCCAGAGAAAGTAACTCTCTCTGTTGGAC	566		
Db	128	CCTGACTCCAAAGGACTCACCGAGGGGAGCTCCAGAGAAAGTAACTCTCTCTGTTGGAC	187		
QY	567	GGGGCTCTGAGAGCTGTCACAGTTTGCATGGCTCTGCTGACAGATCGCTACCTCTCTCT	626		

Db 660 TTAACATTCTGTCATCCACGGCTCATTAATTTTGTGATGATGTTTGG 719
Qy 1078 CGATCCCAATCTGAGGTTTCAAGAGCAAAATGGGTTCTTGGCGCAAGCGACA 1137
Db 720 CGATCCCAATCTGAGGTTTCAAGAGCAAAATGGGTTCTTGGCGCAAGCGACA 779
Qy 1138 CCAATGTCAGAAATGCGGCTATATGATCCGATTCGCCCAACATGGGGTCACT 1197
Db 780 CCAATGTCAGAAATGCGGCTATATGATCCGATTCGCCCAACATGGGGTCACT 839
Qy 1198 GGATCGCCAGAAATCTC-TTTTGAATAATTAATTAAGAACAAAGAGCTTCA 1256
Db 840 GGATCGCCAGAAATCTCTTTTGGAAAAATTAACCTTAGAACAAGAGCTTCA 899
Qy 1257 AATAACAATGGGAGATGGTG 1278
Db 900 AATAACAATGGGAGATGGTG 921

RESULT 4
LOCUS B1912693 930 bp mRNA linear EST 16-OCT-2001
DEFINITION 603176344F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240506 5',
mRNA sequence.
ACCESSION B1912693
VERSION B1912693.1 GI:16176907
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 930)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM11606 row: h column: 11
High quality sequence stop: 810.
Location/Qualifiers
1..930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5240506"
/lab_host="DH10B"
/clone_1fb="NIH_MGC_121"
/clone_2fb="NIH_MGC_121"
/note="Torgan: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 25.8%; Score 823; DB 3; Length 930;
Best Local Similarity 95.7%; Pred. No. 6.2e-189;
Matches 889; Conservative 0; Mismatches 35; Indels 5; Gaps 4;
Qy 279 GCGAGTGTTCAGGTTCTAGAGCTATGACGTGACACTGCTTCTCTCTATCATG 338

Db 1 GCGAGTGTTCAGGTTCTAGAGCTATGACACTGAGAGACTGCTTCTCTCTATCATG 60
Qy 339 CTCCTCCAGAAATTTCTCAATGTGAGAGACGCTACCCATTCAGCGGATCCGAGCT 398
Db 61 CTCCTCCAGAAATTTCTCAATGTGAGAGACGCTACCCATTCAGCGGATCCGAGCT 120
Qy 399 GTCTTGACAGATCATCTCCATTTATCTGACCACTGACCAACTGAGAGAAATTCCTTTG 458
Db 121 GTCTTGACAGATCATCTCCATTTATCTGACCACTGACCAACTGAGAGAAATTCCTTTG 180
Qy 459 AAAAAATTAACAGAGGAGATGACGAATCAGTCAATACAGCAATTTTCTGACTCCAG 518
Db 181 AAAAAATTAACA-GGGAGATGACGAATCAGTCAATACAGCAATTTTCTGACTCCAG 239
Qy 519 GACTCACCAGGGAGCGTCAGAGAAATCTCTCTGCTTTGACGGGGTCTTGAG 578
Db 240 GACTCACCAGGGAGCGTCAGAGAAATCTCTCTGCTTTGACGGGGTCTTGAG 299
Qy 579 CTTCGTCAAGTTTCAGATGCTGTGCTGACAGATCGCTACCTCTCTCAGTCCAGCCAG 638
Db 300 CTTCGTCAAGTTTCAGATGCTGTGCTGACAGATCGCTACCTCTCTCAGTCCAGCCAG 359
Qy 639 CCACAGTCTGCGGCTGCTCTCCAGTGCATGTTCCCGTACCCCGCCACAGCAGACCG 698
Db 360 CCACAGTCTGCGGCTGCTCTCCAGTGCATGTTCCCGTACCCCGCCACAGCAGACCG 419
Qy 699 GCGACACCCCGCTTCTCATTCGAGCGCTTACCGGCTTACGAGTCCGACACCGGGTATC 758
Db 420 GCGACACCCCGCTTCTCATTCGAGCGCTTACCGGCTTACGAGTCCGACACCGGGTATC 478
Qy 759 ACCAAGGAGGCTTCAACAGGCTCTGTCCAACTCTGCGCGAGGATACCCCAAGGCC 818
Db 479 ACCAAGGAGGCTTCAACAGGCTCTGTCCAACTCTGCGCGAGGATACCCCAAGGCC 538
Qy 819 GGCTACCCCTTACCCACAGCAGTACGCGCATCTCTTACCAAGAGCTCCGTTTACAGTTC 878
Db 539 GGCTACCCCTTACCCACAGCAGTACGCGCATCTCTTACCAAGAGCTCCGTTTACAGTTC 598
Qy 879 TCCCTCCACCCGCGCGGGCTGTGTCGCGGCAACACAGGATGATCTGTGACAGGCC 938
Db 599 TCCCTCCACCCGCGGGCTGTGTCGCGGCAACACAGGATGATCTGTGACAGGCC 658
Qy 939 CTTCGCTGAATTTTCAACCGGCAACCAAGGAGATGATCATCACCAACAGGAGAGCGC 998
Db 659 CTTCGCTGAATTTTCAACCGGCAACCAAGGAGATGATCATCACCAACAGGAGAGCGC 718
Qy 999 ATGTTTCTTTTAAAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1058
Db 719 ATGTTTCTTTTAAAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 778
Qy 1059 TTTTGATATGATTTTGGCGGATCCCAATCAGTGAAGTTTCAAGGAGCAATGGGTT 1118
Db 779 TGTGGAATGTGAATGTGGCGGATCCCAATCAGTGAAG-TTCAGAAAGGCAATGGGCT 837
Qy 1119 CTTGCGCGCAAGGCGACCAATGTGCAAGAAATCGGATCTATATGATCCGATTTCC 1178
Db 838 TCTGCGCGCAAGGCGACCAATGTGCAAGAAATCGGATCTATATGATCCGATTTCC 897
Qy 1179 CCACACTGGGGCTCACTGATGGCCA 1207
Db 898 CC-AAATGGGGCTCACTGATGGCCA 924

RESULT 5
LOCUS Bx418656 824 bp mRNA linear EST 03-MAY-2004
DEFINITION Bx418656 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSDDF008Y10 5-PRIME, mRNA sequence.
ACCESSION Bx418656
VERSION Bx418656.2 GI:46954137
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 824)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30765811.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1851.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSODF08BD05QPic=1851.f.
FEATURES
source
1..824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF08YGI0"
/issue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_1ib="Homo sapiens FETAL BRAIN"
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 25.5%; Score 811.4; DB 5; Length 824;
Best Local Similarity 99.8%; Pred. No. 3.9e-185;
Matches 823; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 200 GAAGTCTTTCTTCTAGTGAAGGGGCTGTGATTTCTAATGATTAAGAAGACT 259
DB 1 GATGTCCTTTCTCTAGTGAAGGGGCTGTGATTTCTAATGATTAAGAAGACT 60
QY 260 TAAAAACAGGAGCGGAGGAGGAGTGTGAGTTCTAGAGCTATCAGCTGAGACATG 319
DB 61 TAAAAACAGGAGCGGAGGAGGAGTGTGAGTTCTAGAGCTATCAGCTGAGACATG 120
QY 320 CCTTCTCTTCTATCATGCTCTCAAGAAATTTCTCAATGTAGAGACAGCTACCCACA 379
DB 121 CCTTCTCTTCTATCATGCTCTCAAGAAATTTCTCAATGTAGAGACAGCTACCCACA 180
QY 380 TTGAGGGGAGTCCGAGCTTGTCTTGACCATATCCCATATTTCCACCACTAACACT 439
DB 181 TTGAGGGGAGTCCGAGCTTGTCTTGACCATATCCCATATTTCCACCACTAACACT 240
QY 440 GGAGAGAGTTCACTTTGAAAAAATTTACAGGGGAGTACGAATCAGTCAGATACAGA 499
DB 241 GGAGAGAGTTCACTTTGAAAAAATTTACAGGGGAGTACGAATCAGTCAGATACAGA 300
QY 500 CAATTTTCTGACTCCAAAGACTCACAGGGGAGTGTCAAGAGAATAACTCTCTCTGT 559
DB 301 CAATTTTCTGACTCCAAAGACTCACAGGGGAGTGTCAAGAGAATAACTCTCTCTGT 360
QY 560 CTGGAAGGGGCTCTGAGCTTGTCTCAAGTTTGAAGGCTCTGCTGCAATGCTAAGT 619
DB 361 CTGGAAGGGGCTCTGAGCTTGTCTCAAGTTTGAAGGCTCTGCTGCAATGCTAAGT 420
QY 620 CCTCTCAGTCCAGGAGCGAGCTGTGGGGCACTGCTCCAGTSCCATGTTCCCGTA 679
DB 421 CCTCTCAGTCCAGGAGCGAGCTGTGGGGCACTGCTCCAGTSCCATGTTCCCGTA 480

QY 680 CCCGCGCAGCAGGAGCGAGCGACCCCGCTTCTCCATCGGACCTTACCGCTACAT 739
DB 481 CCCGCGCAGCAGGAGCGAGCGACCCCGCTTCTCCATCGGAGCTTACCGCTACAT 540
QY 740 GGGCCACACCGCGGTATACCAACGAGGCTTACAGCTCTGTTCAACTCTCGCC 799
DB 541 GGGCCACACCGCGGTATACCAACGAGGCTTACAGCTCTGTTCAACTCTCGCC 600
QY 800 GCAAGGATACCCCAACCGCGGCTACCCCTTACCAAGAGTACGAGGCACTCTCAAGG 859
DB 601 GCAAGGATACCCCAACCGCGGCTACCCCTTACCAAGAGTACGAGGCACTCTCAAGG 659
QY 860 AGCTCCGTTTACCAAGTCTCTCTCCACCAAGCGGAGTGTGCTCCGCAAGCAGAGT 919
DB 660 AGCTCCGTTTACCAAGTCTCTCTCCACCAAGCGGAGTGTGCTCCGCAAGCAGAGT 719
QY 920 GTACCTGTGCAACAGGCGCTTTGCTGAAATTTTACCGGACCAACGAGATGATCAT 979
DB 720 GTACCTGTGCAACAGGCGCTTTGCTGAAATTTTACCGGACCAACGAGATGATCAT 779
QY 980 CACCAACAGGAGGAGCGCATGTTCTTTTAAAGTTTAAAT 1024
DB 780 CACCAACAGGAGGAGCGCATGTTCTTTTAAAGTTTAAAT 824
RESULT 6
BI915063 820 bp mRNA linear EST 16-OCT-2001
LOCUS 603177258F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241542 5'
DEFINITION mRNA sequence.
BI915063
ACCESSION BI915063.1 GI:16199126
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
1 (bases 1 to 820)
NIH-MGC http://mgi.mcg.mcg.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLM11609 row: C column: 15
High quality sequence stop: 816.
FEATURES
source
1..820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5241542"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_121"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 25.4%; Score 809; DB 3; Length 820;
 Best Local Similarity 99.9%; Pred. No. 1.5e-184;
 Matches 820; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

221 GGGGGTCTGTGATTTCTAGATTATGATTAATAGACCTTTAAAAACAGAGGACGGAGAGG 280
 1 GGGGGTCTGTGATTTCTAGATTATGATTAATAGACCTTTAAAAACAGAGGAGAGG 59

281 CGAGTTCAGAGTTCTAGAGCTATGAGCTGAGACATGCTCTTCTCTCTATCATGCT 340
 60 CGAGTTCAGAGTTCTAGAGCTATGAGCTGAGACATGCTCTTCTCTCTATCATGCT 119

341 CTCGAGAAATTTCTCAATGTAGAGCAGCAGTACCACTTCCAGGCGGATCCGACTTGT 400
 120 CTCGAGAAATTTCTCAATGTAGAGCAGCAGTACCACTTCCAGGCGGATCCGACTTGT 179

401 CTTCGACGATCATCCCATTTATCTCGACCACTGACAACTGAGAGAGTTCACTTTGAA 460
 180 CTTCGACGATCATCCCATTTATCTCGACCACTGACAACTGAGAGAGTTCACTTTGAA 239

461 AAAAATTCAGAGGGGATGACGATCATGATGACATGATTTCTCTGATCCAGAGA 520
 240 AAAAATTCAGAGGGGATGACGATCATGATGACATGATTTCTCTGATCCAGAGA 299

521 CTCACGAGGGGAGCTCCAGAGAAATTAATCTCTCTCTCTGAGACGGGGTCTGAGGT 580
 300 CTCACGAGGGGAGCTCCAGAGAAATTAATCTCTCTCTCTGAGACGGGGTCTGAGGT 359

581 TCGTCAAGATTTGATGAGTCTGCTGCTGAGATGCTTACCTCTCTCTCACTCAAGCCAGCC 640
 360 TCGTCAAGATTTGATGAGTCTGCTGCTGAGATGCTTACCTCTCTCTCACTCAAGCCAGCC 419

641 AAGAGTCGGGGGACATGCTCCAGAGGACATGTTCCCTGACCCCGGACAGAGACCGGC 700
 420 AAGAGTCGGGGGACATGCTCCAGAGGACATGTTCCCTGACCCCGGACAGAGACCGGC 479

701 GCACCCCGCTTCTCCATGAGGACGCTTACGCTTACATGAGCCCAACCCCGTATCATC 760
 480 GCACCCCGCTTCTCCATGAGGACGCTTACGCTTACATGAGCCCAACCCCGTATCATC 539

761 CAACGAGGCTTCAACAGGCTCTCTGCTCAACTCTCTGCGGAGAGGATACCCACGAGCGG 820
 540 CAACGAGGCTTCAACAGGCTCTCTGCTCAACTCTCTGCGGAGAGGATACCCACGAGCGG 599

821 CTAACCCCTTACCAAGCAGATGAGGCTCACTCTCAACAGAGAGTCCGTTTACAGATTC 880
 600 CTAACCCCTTACCAAGCAGATGAGGCTCACTCTCAACAGAGAGTCCGTTTACAGATTC 659

881 CTCACCCAGCGGGGCTGTGCTCCGCGCAAGACAGAGTGTACTGTGCAACAGGCGCT 940
 660 CTCACCCAGCGGGGCTGTGCTCCGCGCAAGACAGAGTGTACTGTGCAACAGGCGCT 719

941 TTGGCTGAATTTTACCGGACCAACGAGAGATGATCATACCAACAGGAGAGGCGCAT 1000
 720 TTGGCTGAATTTTACCGGACCAACGAGAGATGATCATACCAACAGGAGAGGCGCAT 779

1001 GTTTCCTTTTAAAGTTTAAATTTCTGCTCTGATCCCA 1041
 780 GTTTCCTTTTAAAGTTTAAATTTCTGCTCTGATCCCA 820

RESULT 7
 BI914330 872 bp mRNA 1linear EST 17-OCT-2001
 LOCUS BI914330
 DEFINITION 60318231fl NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246210 5',
 mRNA sequence.
 ACCESSION BI914330
 VERSION BI914330.1 GI:16198838
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE
 1 (bases 1 to 872)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1AM1621 row: F column: 03
 High quality sequence stop: 853.

FEATURES
 source
 1..872
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246210"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_121"
 /note="Organ: brain; Vector: pCMV-Sport6; site_1: NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 25.2%; Score 801.6; DB 3; Length 872;
 Best Local Similarity 97.7%; Pred. No. 9.4e-183;
 Matches 856; Conservative 0; Mismatches 14; Indels 6; Gaps 4;

367 GCAGTACCCATTCAGAGCGGATCCGAGCTTGTCTTGACAGATCATCCATTATCTCGA 426
 1 GCAGTACCCATTCAGAGCGGATCCGAGCTTGTCTTGACAGATCATCCATTATCTCGA 60

427 CCACGACACCTGGAGAGAGTTCACTTTGAAAAAATTCACAGGGGATGACGAATC 486
 61 CCACGACACCTGGAGAGAGTTCACTTTGAAAAA--TTACAGGGGATGACGAATC 118

487 AGTCAGATACAGACATTTTCTGATCTCAAGAGCTCACAGGGAGAGTCCAGAGAAATA 546
 119 AGTCAGATACAGACATTTTCTGATCTCAAGAGCTCACAGGGAGAGTCCAGAGAAATA 178

547 AACTCTCTCTGTCTTGAGCGGGGTCTGTAGCTTCTGACAGTTTGCATGAGTCTGTG 606
 179 AACTCTCTCTGTCTTGAGCGGGGTCTGTAGCTTCTGACAGTTTGCATGAGTCTGTG 238

607 CAGATGGTACTCTCTCTGATGACGCGACAGCAGTGTGCGGACCTGCTCCAGTG 666
 239 CAGATGGTACTCTCTCTGATGACGCGACAGCAGTGTGCGGACCTGCTCCAGTG 298

667 CCATGTTCCCGTACCCCGGCGAGCAGCAGCCGCGGCAACCCGCTTCTCATGCGAGCC 726
 299 CCATGTTCCCGTACCCCGGCGAGCAGCAGCCGCGGCAACCCGCTTCTCATGCGAGCC 358

727 CTAGCGCTTACATGAGCCACACCGGGTATCATCACCAAGAGCTTCAACAGCTCTGT 786
 359 CTAGCGCTTACATGAGCCACACCGGGTATCATCACCAAGAGCTTCAACAGCTCTGT 418

787 CCAACTCTCTGCGCGAGGGATACCCACAGGCGGCTACCTTACCCACAGAGTACGAGCC 846
 419 CCAACTCTCTGCGCGAGGGATACCCACAGGCGGCTACCTTACCCACAGAGTACGAGCC 478

847 ACTCTTACCAAGAGCTCGTTTACAGATTCTCTTCAACCCAGCCGCGGCTGTGCGCG 906

Db 479 ACTCTACCAAGAGAGTCCGTTCTACCAATCTCTCCACCAACCGGCGGTGTCGG 538
 Qy 907 GCMAAGCAGAGGTGTAACCTGTGCAACAGGCCCTTTGGCTGAATTTCAACGGCAGCAAA 966
 Db 539 GCMAAGCAGAGGTGTAACCTGTGCAACAGGCCCTTTGGCTGAATTTCAACGGCAGCAAA 598
 Qy 967 CGGAGATGATCATCACCAACAGGGAGCGCATGTTCTCTTTTAAGTTTAACATTT 1026
 Db 599 CGGAGATGATCATCACCAACAGGGAGCGCATGTTCTCTTTTAAGTTTAACATTT 658
 Qy 1027 CTGGCTGATCCGACGCTCATTTCAATTTTGTGATGATGATTTGGCGATCCCA 1086
 Db 659 CTGGCTGATCCGACGCTCATTTCAATTTTGTGATGATGATTTGGCGATCCCA 718
 Qy 1087 ATCACTGAGAGTTTCAAGAGGCAATGGGTTCTTGGCGCAAGCGGACCAATGTGC 1146
 Db 719 ATCACTGAGAGTTTCAAGAGGCAATGGGTTCTTGGCGCAAGCGGACCAATGTGC 778
 Qy 1147 AAGGAAA--TCGGGTCTATATGATCGGATTCGCCCAACACTGGGGCTCACTGATGCG 1204
 Db 779 AAGGAAAATCGGGGTCTATATGATCGGATTCGCCCAACACT--GGGCTCACTGATGCG 837
 Qy 1205 CCAAGAAATCTCTTTGGAAAATTAATTAAGTTACGA 1240
 Db 838 -CAAGAAATCTCTTTGGAAAATTAATTAAGTTACGA 872

RESULT 8
 LOCUS B0067178 1169 bp mRNA linear EST 02-APR-2002
 DEFINITION AGENCOURT_6767481 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5751547
 5', mRNA sequence.
 B0067178
 VERSION B0067178.1 GI:19896224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 1169)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12784 row: e column: 20
 High quality sequence start: 16
 High quality sequence stop: 685.
 Location/Qualifiers
 1..1169
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5751547"
 /lab_host="DH10B"
 /note="Organ: pooled brain; lung; testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.

FEATURES

source

Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."
 Query Match 25.1%; Score 798.2; DB 3; Length 1169;
 Best Local Similarity 96.4%; Pred. No. 6.7e-182;
 Matches 826; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

Qy 826 CCTACCCACAGCATGACCGCCATCTTACCAAGAGAGTCCGTTCTCAACAGTTTCCTCA 885
 Db 36 CCTACCCACAGCATGACCGCCATCTTACCAAGAGAGTCCGTTCTCAACAGTTTCCTCA 95
 Qy 886 CCGAGCCGGGCTGGTGCCTGGCAAGAGCAGTGTACTGTGCAACAGGCCCTTTGGC 945
 Db 96 CCGAGCCGGGCTGGTGCCTGGCAAGAGCAGTGTACTGTGCAACAGGCCCTTTGGC 155
 Qy 946 TGAATTTTACCGGACCAACCGAGATGATCATCACCAACAGGGAAGCGCATGTTTC 1005
 Db 156 TGAATTTTACCGGACCAACCGAGATGATCATCACCAACAGGGAAGCGCATGTTTC 215
 Qy 1006 CTTTTTAAGTTTAACTTTCTGTGATCCCAAGGCTCATTAACAATTTTGTG 1065
 Db 216 CTTTTTAAGTTTAACTTTCTGTGATCCCAAGGCTCATTAACAATTTTGTG 275
 Qy 1066 ATGTGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGGTTCCTTGC 1125
 Db 276 ATGTGATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGGCAATGGTTCCTTGC 335
 Qy 1126 GCMAAGCAGACCAACATGTGCAAGGAAATCGGGTCTATATGATCGGATTCGCCCAACA 1185
 Db 336 GCMAAGCAGACCAACATGTGCAAGGAAATCGGGTCTATATGATCGGATTCGCCCAACA 395
 Qy 1186 CTGGGGCTCACTGATGCGCCCAAGAAATCTTTTGGAAAATTAACCTTACGACACA 1245
 Db 396 CTGGGGCTCACTGATGCGCCCAAGAAATCTTTTGGAAAATTAACCTTACGACACA 455
 Qy 1246 AAGGAGCTTCAAAATAACATGGGCGATGTGTTTACAGTCTTGGACAAGTACGAC 1305
 Db 456 AAGGAGCTTCAAAATAACATGGGCGATGTGTTTACAGTCTTGGACAAGTACGAC 515
 Qy 1306 CCGGCTGATGTGTGGAAGTGAACGAGAGCGGAGGACATGACCGCCGCGC 1365
 Db 516 CCGGCTGATGTGTGGAAGTGAACGAGAGCGGAGGACATGACCGCCGCGC 575
 Qy 1366 GCGTGAACGTTCACTTTCTTCTGAGACTCACTTATGCGCGTACCGGCTTACGAAACA 1425
 Db 576 GCGTGAACGTTCACTTTCTTCTGAGACTCACTTATGCGCGTACCGGCTTACGAAACA 635
 Qy 1426 CGGATATTACCACTGAATAATGATCAACCCCTTTGCAAAAGATTCGGATTAAT 1485
 Db 636 CGGATATTACCACTGAATAATGATCAACCCCTTTGCAAAAGATTCGGATTAAT 695
 Qy 1486 ATGACAGATCTACACCGGCTGTGATGACCGGCTGACCCCTTGGCCCAACGACTGCG 1545
 Db 696 ATGACAGATCTACACCGGCTGTGATGACCGGCTGACCCCTTGGCCCAACGACTGCG 755
 Qy 1546 CCGCTCGCAGATGCTGCCCGGGGCCGCTACCGCATGGCGGCTTTCTTCTGCA-GGAC 1604
 Db 756 CCGCTCGCAGATGCTGCCCGGGGCCGCTACCGCATGGCGGCTTTCTTCTGCAAGGAC 815
 Qy 1605 CAGTTCGTGAGCACTACGCGCAAGGCGCGCTTCCACCGCGGCGCGGCGCGCGG 1664
 Db 816 CAGTTCGTGAGCACTACGCGCAAGGCGCGCTTCCACCGCGGCGCGGCGCGGCGG 875
 Qy 1665 CCGGCTACGACCGGAG 1681
 Db 876 CCGGCTACGACCGGAG 892

RESULT 9
 LOCUS BM922065 1044 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT_6707015 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5754122

5', mRNA sequence.
 BM922065
 BM922065.1 GI:19372444
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L16M12791 row: a column: 03
 High quality sequence stop: 694.
 Location/Qualifiers
 1..1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5754122"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC library."

ORIGIN

Query Match 24.8%; Score 789.8; DB 3; Length 1044;
 Best Local Similarity 97.5%; Pred. No. 7.1e-180;
 Matches 866; Conservative 0; Mismatches 13; Indels 9; Gaps 6;

202 AGTGCCTTCTGTCTGTGAGGGGCTGTGAGTTTCTAGTTATGATTAATGAGCTTTA 261
 |||||
 23 AGTGCCTTCTGTCTGTGAGGGGCTGTGAGTTTCTAGTTATGATTAATGAGCTTTA 82
 |||||
 262 AAAACGAGGAGCGGAGGCGAGTGTTCAGTTCCTAGAGCTATGAGCTGAGCACTGCC 321
 |||||
 83 AAAACGAGGAGCGGAGGCGAGTGTTCAGTTCCTAGAGCTATGAGCTGAGCACTGCC 142
 |||||
 322 TTTCTCTTCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGACCTTACCACTT 381
 |||||
 143 TTTCTCTTCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGACCTTACCACTT 202
 |||||
 382 CAGGCGAGATCCAGAGCTGTCTGACAGATATCCCATTTCTGACCACTGACCACTGG 441
 |||||
 203 CAGGCGAGATCCAGAGCTGTCTGACAGATATCCCATTTCTGACCACTGACCACTGG 262
 |||||
 442 AGAGAGTTACCTTTGAATAAAATTTACAGGCGGAGTGAAGATCAGTCAATACAGACA 501
 |||||
 263 AGAGAGTTACCTTTGAATAAAATTTACAGGCGGAGTGAAGATCAGTCAATACAGACA 322
 |||||
 502 ATTTTCTGACTTCAAGACTCAACGAGGAGCTCCAGAGAAATTAATCTCTCTGTTCT 561
 |||||
 323 ATTTTCTGACTTCAAGACTCAACGAGGAGCTCCAGAGAAATTAATCTCTCTGTTCT 382
 |||||
 562 TGAGCGGGGTCTCTGAGCTTCTGCAAGTTTGCATGGCTCTGCTGAGATCCGTAACCTCC 621

383 TGAGCGGGGTCTCTGAGCTTCTGTCACAGTTTGCATGGCTCTGCTGAGATCGTACTCC 442
 |||||
 622 TCTCTCAGTCCAGGCAAGCCACAGTCTGCGGCACTGCTCCAGTGCATGTTCCGTAAC 681
 |||||
 443 TCTCTCAGTCCAGGCAAGCCACAGTCTGCGGCACTGCTCCAGTGCATGTTCCGTAAC 502
 |||||
 682 CCGGCGAGCAAGGAGCGGCGGACCCGCTTCTGATCGGAGCGCTTACGCGCTACATGG 741
 |||||
 503 CCGGCGAGCAAGGAGCGGCGGACCCGCTTCTGATCGGAGCGCTTACGCGCTACATGG 562
 |||||
 742 CCCACACCGCGTATCATCAACAGGAGCTTACCAACAGCTTCTGCTCACTCTCGCGCG 801
 |||||
 563 CCCACACCGCGTATCATCAACAGGAGCTTACCAACAGCTTCTGCTCACTCTCGCGCG 622
 |||||
 802 AGGATATCCCAACGCGCGGCTACCTCTTACCAACAGCTTACGCTTCTTCCAGAGAG 861
 |||||
 623 AGGATATCCCAACGCGCGGCTACCTCTTACCAACAGCTTACGCTTCTTCCAGAGAG 682
 |||||
 862 CTCGGTCTTACCACTTCTCTCCACCGAGCGGCGGCTGTCGCGGCAAGCAAGGTGT 921
 |||||
 683 CTCGGTCTTACCACTTCTCTCCACCGAGCGGCGGCTGTCGCGGCAAGCAAGGTGT 742
 |||||
 922 ACTGTGCAACAGGCGCCCTTGTGCTGAATTTTACCGGCGCAACCAAGAGATGATCAT 979
 |||||
 743 ACTGTGCAACAGGCGCCCTTGTGCTGAATTTTACCGGCGCAACCAAGAGATGATCAT 802
 |||||
 980 CACCAACA-GGAGAGGCGAGTGTTCCTT-TTTTAAAGTTTAAACATTTCTGCTCTC--- 1034
 |||||
 803 CACCAACAAGAGAGGCGAGTGTTCCTT-TTTTAAAGTTTAAACATTTCTGCTCTC--- 1034
 |||||
 1035 GATCCACGCGCTCATTAATTTT--GTGATGTGATTTTGGCGG 1080
 |||||
 863 ATCCCAAGGCTCATTAATTTTGTGGATGTGATTTGGGG 910
 |||||

RESULT 10
 B1914132
 LOCUS B1914132 945 bp mRNA linear EST 16-OCT-2001
 DEFINITION 603182077F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246298 5',
 mRNA sequence.
 B1914132
 ACCESSION B1914132.1 GI:16178444
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L16M11621 row: i column: 19
 High quality sequence stop: 711.
 Location/Qualifiers
 1..945
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5246298"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;

FEATURES
 source

ORIGIN

Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."

Query Match 24.1%; Score 766.6; DB 3; Length 945;
Best Local Similarity 92.0%; Pred. No. 2.9e-174;
Matches 866; Conservative 0; Mismatches 64; Indels 11; Gaps 5;

```

QY 282 GAGTTTCAGGTTCTAGAGCTATGACAGTGAAGCACTGCTTCTCTTCTATATGCTC 341
DB 2 GCGAGTGTAGGTTCTAGAGCTATGACAGTGAAGCACTGCTTCTCTTCTATATGCTC 61
QY 342 TCCAGAAATTTCTCAATGTGAGCAGCAGCTTACCAATTCAGCGGATCCGAGCTTTC 401
DB 62 TCCAGAAATTTCTCAATGTGAGCAGCAGCTTACCAATTCAGCGGATCCGAGCTTTC 121
QY 402 TTGCAAGATCATCTCCATTTCTCGACCTGCAACCTGAGAGAAAGTTCACTTTGAAA 461
DB 122 TTGCAAGATCATCTCCATTTCTCGACCTGCAACCTGAGAGAAAGTTCACTTTGAAA 181
QY 462 AAAATATCCAGGGGGGATGAGCAATCATGATACAGACAAATTTCTGACTCCAGAGAC 521
DB 182 AAAATATCCAGGGGGGATGAGCAATCATGATACAGACAAATTTCTGACTCCAGAGAC 241
QY 522 TCACAGAGGAGAGCTCCAGAGAAATATCTCTCTCTTGGACGGGCTCTTGAGCTT 581
DB 242 TCACAGAGGAGAGCTCCAGAGAAATATCTCTCTCTTGGACGGGCTCTTGAGCTT 301
QY 582 CGTCAAGATTTGATAGGCTTCTGCTGAGATCGGTACTCTCTCTGATTCAGCCAGCA 641
DB 302 CGTCAAGATTTGATAGGCTTCTGCTGAGATCGGTACTCTCTCTGATTCAGCCAGCA 361
QY 642 CAGTCTGCGGCACTGCTCCAGATGCAATGTTCCCGTACCCCGGAGAGACAGGAGCG 701
DB 362 CAGTCTGCGGCACTGCTCCAGATGCAATGTTCCCGTACCCCGGAGAGACAGGAGCG 421
QY 702 CACCCGCGCTTCTCAATCGGAGCCCTTACGCGGTATAGCCACCAACCGGTCATCACC 761
DB 422 CACCCGCGCTTCTCAATCGGAGCCCTTACGCGGTATAGCCACCAACCGGTCATCACC 481
QY 762 AAGGAGCTTCAACAGCTCTCTGCTCAATCTCTGCGCGGAGGATACCCAGGCGGCG 821
DB 482 AAGGAGCTTCAACAGCTCTCTGCTCAATCTCTGCGCGGAGGATACCCAGGCGGCG 541
QY 822 TACCCCTACCCAGAGAGTACGAGGCACTCTTACCAAGAGAGCTCCGTTCTACAGTTCTCC 881
DB 542 TACCCCTACCCAGAGAGTACGAGGCACTCTTACCAAGAGAGCTCCGTTCTACAGTTCTCC 601
QY 882 TCCACCCAGCGGCGGCTGATGCGCGGCAAGCAAGGTGTAAGTGTGCAACAGGCGCTT 941
DB 602 TCCACCCAGCGGCGGCTGATGCGCGGCAAGCAAGGTGTAAGTGTGCAACAGGCGCTT 661
QY 942 TGGCTGAATTTTCAACCGGCAACAA-AGAGATGATATATACCAAAAGGAGAGGCGCAT 1000
DB 662 TGGCTGAATTTTCAACCGGCAACAAAGGAGATATATACCAAAAGGAGAGGCGCAT 721
QY 1001 GTTTCCTTTT-----AAGTTTAAATTTCTGG--TCTGATCCCAAGGCTCATTTACA 1054
DB 722 TGTTCCTTTTGAAGTTTAAATTTCTGGGTCTGGAACCCCAAGGCTCATTTACA 781
QY 1055 TATTTTGTGAGTGTATTTTGGGAGATCCCAATCATGAGGTTTCAAGAGGCAATG 1114
DB 782 TATTTGTGAGTGTATTTGGGAGATCCCAATCATGAGGTTTCAAGAGGCAATG 841
QY 1115 GGTTCCTTGGGAGAGAGCAACCAATGTGCAAGAAATCGGGTCTATATGATCCGGA 1174

```

```

DB 842 GTTCTGACAGGCAAAACGACACCCAT-TGTGAGGAATTCGGGCTATATGATCCGAT 900
QY 1175 TTCCCCAACACCTGGGGCTACCTGATGCGGCAAGAAATCT 1215
DB 901 TCCCCAACAC--TGGGCTATGGGATGCGCAAGAAATCTCT 938

```

RESULT 11

B1754056 966 bp mRNA linear EST 25-SEP-2001
LOCUS 60302761J1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5197826 5'

DEFINITION RNA sequence.

ACCESSION B1754056

VERSION B1754056.1 GI:15745634

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.

REFERENCE 1 (bases 1 to 966)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LMNL at:

<http://image.llnl.gov>

Plate: LHM11495 row: F column: 03

High quality sequence stop: 816.

FEATURES

source

1..966

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5197826"

/lab_host="DH10B"

/note="Torgan: brain; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 23.9%; Score 761; DB 3; Length 966;
Best Local Similarity 92.2%; Pred. No. 6.7e-173;
Matches 904; Conservative 0; Mismatches 60; Indels 17; Gaps 9;

```

QY 203 GTGCTTCTGTCTAGT-AGGGGATCTGTGATTTCTAGTTATGATTAATGACTTTA 261
DB 1 GTGCTTCTGTCTAGTAAAGGGGCTGTGATTTCTAGTTATGATTAATGACTTTA 60
QY 262 AAAACAGGAGAGGAGGAGGAGTTCAGGTTCTAGAGCTTATGAGAGCTGAGCACTGCC 321
DB 61 AAAACCA-AGAGGAGAGGAGGAGGAGTTCAGGTTCTAGAGCTTATGAGAGCTGAGCACTGCC 119
QY 322 TTTCTCTTCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGACTACCAAT 381
DB 120 TTTCTCTTCTATCATGCTCTCCAGAAATTTCTCAATGTGAGCAGACTACCAAT 179
QY 382 CAGGCGATCCAGAGCTTGTCTTGAAGATCCATATATCTGACCACTGACCAACTGG 441
DB 180 CAGGCGATCCAGAGCTTGTCTTGAAGATCCATATATCTGACCACTGACCAACTGG 239

```

QY 442 AGAGAAGTTCACTTGGAAAAAATTTACAGGGGGATGACGATCATGTCAGATACAGCA 501
 DB 240 AGAGAAGTTCACTTGGAAAAAATTTACAGGGGGATGACGATCATGTCAGATACAGCA 299
 QY 502 ATTTTCCTGACTCCAAAGACTCAGCAGGGGAGCGTCCAGAGAGTAAGTAACTCTCTCTCTCT 561
 DB 300 ATTTTCCTGACTCCAAAGACTCAGCAGGGGAGCGTCCAGAGAGTAAGTAACTCTCTCTCTCT 359
 QY 562 TGGACGGGGTCTCTGAGCTTGTGTCACAGTTTGATGAGCTCTGCTGAGATCGCTACCTCC 621
 DB 360 TGGACGGGGTCTCTGAGCTTGTGTCACAGTTTGATGAGCTCTGCTGAGATCGCTACCTCC 419
 QY 622 TCTTCAGTCCAGCAGCAGCAGCTGTCGGGGGCACTGCTCCAGTGCATGTTCCCGTACC 681
 DB 420 TCTTCAGTCCAGCAGCAGCAGCTGTCGGGGGCACTGCTCCAGTGCATGTTCCCGTACC 479
 QY 682 CCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 741
 DB 480 CCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 539
 QY 742 CCCACACCCGGTCACTACCAACGAGCTTACAGAGCTCTCTGTCACATCTCTGCGCGC 801
 DB 540 CCCACACCCGGTCACTACCAACGAGCTTACAGAGCTCTCTGTCACATCTCTGCGCGC 599
 QY 802 AGGGAATACCCCAAGCGCGGCTACCCCTACCCAGAGATGAGGCGCACTCTACCAAGGAG 861
 DB 600 AGGGAATACCCCAAGCGCGGCTACCCCTACCCAGAGATGAGGCGCACTCTACCAAGGAG 658
 QY 862 CTCGGTTCACAGTTCCTCTCCACCCAGCAGCGGGGCTGCTGCGCGCAAGCAGAGTGT 921
 DB 659 CTCGGTTCACAGTTCCTCTCCACCCAGCAGCGGGGCTGCTGCGCGCAAGCAGAGTGT 718
 QY 922 ACCGTGCAACAGCGCCCTTTGGCTGAATTTCAACCGGCAACAAACGAGATGATCATCA 981
 DB 719 ACCGTGCAACAGCGCCCTTTGGCTGAATTTCAACCGGCAACAAACGAGATGATCATCA 776
 QY 982 CCAACAGGGAGAGCGCATGTTCTCTTTTAAGTTTAAATTTGATGCTCGATCCCA 1041
 DB 777 CCAACAGGGAGAGCGCATGTTCTCTTTTAAGTTTAAATTTGATGCTCGATCCCA 834
 QY 1042 CGGCTCATTAATATTTTGTGATGATGATTTGGCGGATCCCAATCACTGAGAGTTTC 1101
 DB 835 GGTTCATACATA---TTGCGTGGCTGTGATTTGGGGGG--TCCATCCCTGCGGGGTTTC 888
 QY 1102 AAGGAGGCAATAGGTTCTCTGCGGCAAGCGGACCAATGTGCAAGGAATCGGGTCT 1161
 DB 889 AAGGAGGCAATAGGTTCTCTGCGGAG--AAGCGGACCCCAATTTGCC--GGGAACCGGGTCT 944
 QY 1162 ATATGATCCGGAATTCGCCCA 1182
 DB 945 ATTTGCGTCCCGTTCCCA 965
 RESULT 12
 AM003023/c 831 bp mRNA linear EST 08-MAR-2000
 LOCUS AM003023/wg61d04.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2475751.3
 DEFINITION similar to SW:TBRI_HUMAN Q16650 T-BRAIN-1 PROTEIN; contains PTRS.t3
 TARI repetitive element ;, mRNA sequence.
 ACCESSION AM003023
 VERSION AM003023.1 GI:5849861
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 831)
 NCI-GAP <http://www.ncbi.nlm.nih.gov/ncigap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrip/image/image.html
 Insert Length: 880 Std Error: 0.00
 Seq primer: -40UP from Gldco
 High quality sequence stop: 458.
 location/Qualifiers
 1..831
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2475751"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_id="NCI CGAP GC6"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
 from the normalized library NCI CGAP GC4 was prepared, and
 ss circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (cloneids
 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 23.9%; Score 760; DB 1; Length 831;
 Best Local Similarity 94.8%; Pred. No. 1,16-172;
 Matches 791; Conservative 0; Mismatches 40; Indels 3; Gaps 1;

QY 1146 CAAGGAATGGGTCATATGATCCGAGTCCCAACCTGGGGCTCACTGATGCGC 1205
 DB 831 CGAGGAATGGGTCATATGATCCGAGTCCCAACCTGGGGCTCACTGAA---GC 775
 QY 1206 CAAGGAATCTTTTGGAAAAATTTAACTTACGAACAAGAGGCTTCAATTAACAT 1265
 DB 774 GCAAGGAATCTTTTGGAAAAATTTAACTTACGAACAAAGAGGCTTCAATTAACAAAT 715
 QY 1266 GGGCAGATGGTGTTCATGATCTCTTGCACAAATGACAGCCCGCTGCAATGTGTGAA 1325
 DB 714 GGGCAGATGGTGTTCATGATCTCTTGCACAAATGACAGCAGCTGCAATGTGTGAA 655
 QY 1326 GTGAACGAGGAGCGGACGAGGACCTAGCCAGCCGCGCGCTGCAAGAGCTTCACTTTC 1385
 DB 654 GTGANCGAGACGAGGAGGACCTAGCCAGCCGCGCGCTGCAAGAGCTTCACTTTC 595
 QY 1386 CCGTGAATCACTTATGATCCGCTGACCGCTTACGAGAACAGATATTACAACTGAAA 1445
 DB 594 CCGTGAATCACTTATGATCCGCTGACCGCTTACGAGAACAGATATTACAACTGAAA 535
 QY 1446 ATAGATCAACACCTTTTGGAAAAAGATTTGGGATATTATGACAGATTTACACGGCC 1505
 DB 534 ATAGATCAACACCTTTTGGAAAAAGATTTGGGATATTATGACAGATTTACACGGCC 475
 QY 1506 TGTGACATGAGCGCGCTGACCCCTGGCCAAAGACTGCGCGGCTGCGAGATGTGCGCC 1565
 DB 474 TGTGACATGAGCGCGCTGACCCCTGGCCAAAGACTGCGCGGCTGCGAGATGTGCGCC 415
 QY 1566 GGGGCGCGCTGACGATGCGCGGCTCTTTCTTGGAGGACCAAGTTCGTGAGCACTAGCGCC 1625
 DB 414 GGGGCGCGCTGACGATGCGCGGCTCTTTCTTGGAGGACCAAGTTCGTGAGCACTAGCGCC 355
 QY 1626 AAGGCGCGCTTCAACCGGGGCGCGGGGCGCGGGCGCGGGTACGAGCCGAGCGTG 1685
 DB 354 AAGGCGCGCTTCAACCGGGGCGCGGGGCGCGGGCGCGGGTACGAGCCGAGCGTG 295

QY 1686 CCGACACCAACGAGGCTGCTGTGCTGCGGACAGAGCCGAGAGCCCGGCGCCCTTGGCGG 1745
 Db 294 CCGACACCAACGAGGCTGCTGTGCTGCGGACAGAGCCGAGAGCCCGGCGCCCTTGGCGG 235
 QY 1746 CAACGCTGTTGTGTGACGCGCGGCAACACCGGCTGAGCTTGCGGGGCTCGGCTTATGAC 1805
 Db 234 CAACGCTGTTGTGTGACGCGCGGCAACACCGGCTGAGCTTGCGGGGCTCGGCTTATGAC 175
 QY 1806 ACAGCAGCAGGACTTCCGCGGCAACGCGGCAACGCTGCTCTTATACGCGGCGGCGGCGTG 1865
 Db 174 ACAGCAGCAGGACTTCCGCGGCAACGCGGCAACGCTGCTCTTATACGCGGCGGCGGCGTG 115
 QY 1866 AAGGCGCTGCGCTGCAAGCTGCAAGCTGCACTGCGCGCGGCTGCGGCTTATGACGCGGAC 1925
 Db 114 AAGGCGCTGCGCTGCAAGCTGCAAGCTGCACTGCGCGCGGCTGCGGCTTATGACGCGGAC 55
 QY 1926 CCGTGGGCTGGGGCGCGCGGCACTGCGGCACTGCGGCGGCAAGTGGGCG 1979
 Db 54 CCGTGGGCTGGGGCGCGCGGCACTGCGGCACTGCGGCGGCAAGTGGGCG 1
 RESULT 13
 CN643299 910 bp mRNA linear EST 12-MAY-2004
 LOCUS ILLUMIGN_MCO_7632_Katze_MBR_Macaca mulatta cDNA clone 1810W:3780
 DEFINITION 5' similar to Bases 1 to 810 highly similar to human TBR1 (Hs_210862), mRNA sequence.
 ACCESSION CN643299
 VERSION CN643299.1 GI:47154309
 KEYWORDS EST.
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecoidea; Cercopithecoidea; Macaca.
 1 (bases 1 to 910)
 MAGNES, C.T., FELLIN, P.C., THOMAS, M.J., KORTH, M.J., AGY, M.B., PROLL, S.C., FITZGIBBON, M., SCHERER, C.A., MINER, D.G., KATZE, M.G. and IADONATO, S.P.
 Analysis of the Macaca mulatta transcriptome and the sequence divergence between Macaca and human
 genome Biol. 6 (7), R60 (2005)
 15998449
 JOURNAL Contact: C. Magnes
 PUBMED 11umigen Biosciences Inc.
 COMMENT 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400
 Fax: 2063780408
 Email: cmagnes@11umigen.com
 Sequenced on 2004.01.02. 826 Q20 bases.
 PCR Primers
 FORWARD: CCCTCACTAAGGAGCAACAAA
 BACKWARD: CACTATAGGCGCAATTTGGTA
 Insert Length: 910 Std Error: 0.00
 Plate: CL000059 row: D column: 08
 Seq primer: CCCTCACTAAGGAGCAACAAA
 POLYA=No.
 FEATURES
 source location/Qualifiers
 1. 910
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /clone="IB10W:3780"
 /sex="female"
 /dev_stage="adult"
 /lab_host="E. coli SOLR"
 /clone_1b="Katze_MBR"
 /note="Organ: brain; Vector: Uni-ZAP XR; Site 1: Ecor I; Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis kit (catalog #200400) and ZAP-cDNA Gigapack III Gold Cloning Kit (catalog #200450)"

ORIGIN

Query Match 23.5%; Score 749.6; DB 7; Length 910;
 Best Local Similarity 97.0%; Pred. No. 3.8e-170;
 Matches 786; Conservative 0; Mismatches 19; Indels 5; Gaps 2;
 QY 2382 CCGGACCCCGGACGAGCCCGGACAGGCTCTTCCGAGTCCGCTCCCACTCTCC 2441
 Db 1 CCGGACCCCGGACGAGCCCGGACAGGCTCTTCCGAGTCCGCTCCCACTCTCC 60
 QY 2442 TTGCGACCCCACTATTTATTTGAACCTCGATGCGCGCTGCGACGAAATAGCAGGT 2501
 Db 61 TTGCGACCCCGGCTATTTATTTGAACCTCGATGCGCGCTGCGACGAAATAGCAGGT 120
 QY 2502 CTCGAGCGGTATTTTAAAC-TTTTTGAACAGAGCTCTTGCATTTAGTACCGACCT 2560
 Db 121 CTCGAGCGGTATTTTAAACCTTTTTTGAACAGAGCTCTTGCATTTAGTACCGACCT 180
 QY 2561 TCAACTTTGCTGAACCTTTTGTCTTCTTACTCTTCTTCTGAGATTAATCTC 2620
 Db 181 TCAACTTTGCTGAACCTTTTGTCTTCTTACTCTTCTTCTGAGATTAATCTC 240
 QY 2621 CTACAAATCCCTCCCTCCCTCGATCTTCTTAACTCTTAACTCTTCTTCTTGAATGAA 2680
 Db 241 CTACAAATCCCTCCCTCCCTCGATCTTCTTAACTCTTAACTCTTCTTCTTGAATGAA 300
 QY 2681 CTCTTACCTTTAGAGACCTGCGGACGCTCTGACGAGCGCATTTCCGACCGGCA 2740
 Db 301 CTCTTACCTTTAGAGACCTGCGGACGCTCTGACGAGCGCATTTCCGACCGGCA 360
 QY 2741 GTCGCGGCTGCAATTTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2800
 Db 361 GTCGCGGCTGCAATTTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 2801 TCCCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2860
 Db 421 TCCCTTCTTATCCCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 2861 AGTGAACCAAGCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2920
 Db 481 AGTGAACCAAGCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 2921 GTTTTATCTCTGTAAGGCTCTAGAAACATGAAATCTAGTATGATGATGATGATGAT 2980
 Db 541 GTTTTATCTCTGTAAGGCTCTAGAAACATGAAATCTAGTATGATGATGATGATGAT 600
 QY 2981 CTTTGTCTTCT 3036
 Db 601 CTTTGTCTTCT 660
 QY 3037 GTGAGTAATATTTCTAGAGGCTCTAGAAACATGAAATCTAGTATGATGATGATGAT 3096
 Db 661 GTGAGTAATATTTCTAGAGGCTCTAGAAACATGAAATCTAGTATGATGATGATGAT 720
 QY 3097 ACTTCTCTCTCAATCCGTTGCAAGAAATATTTCTATGTCCTTAATGACACAATAGCT 3156
 Db 721 ACTTCTCTCTCAATCCGTTGCAAGAAATATTTCTATGTCCTTAATGACACAATAGCT 780
 QY 3157 AAGAGAAATCCACCCCAACCTTTAAAG 3186
 Db 781 AAGAGAAATCCACCCCAACCTTTAAAG 810
 RESULT 14
 DR001136 768 bp mRNA linear EST 17-MAY-2005
 LOCUS TC115971_Human fetal brain, large insert, pCMV expression library
 DEFINITION Homo sapiens cDNA clone TC115971 5' similar to Homo sapiens T-box, brain, 1 (TBR1), mRNA sequence.
 ACCESSION DR001136
 VERSION DR001136.1 GI:66261009
 KEYWORDS EST.
 SOURCE Homo sapiens (human)


```

/sex="female"
/dev string="adult"
/lab host="E. coli SOLR"
/clone_lib="Katze_MMR"
/notes="Organ: brain; Vector: Uni-ZAP XR; Site: 1: EcoR I;
Site 2: Xho I; Created from Stragene ZAP-CDNA Synthesis
kit (catalog #200400) and ZAP-CDNA Gigapack III Gold
Cloning Kit (catalog #200450) "

```

ORIGIN

Query Match	Score	DB 7;	Length
22.68;	720;	DB 7;	Length 910;

Best Local Similarity 95.0%; Pred. NO. 5./6-103;
Matches 774; Conservative 0; Mismatches 30; Indels 6; Gaps 3

[illegible]

Search completed: December 20, 2005, 15:56:22
Job time : 12035 secs

This Page Blank (uspto)

Db 241 GACTCACGAGGGAGCGTCCAGAGAAATTAATCTCTCTGTTCTTGGAAGGGGCTCTAG 300
QY 579 CTTGCTACAGTTTGGATGGCTCTGTGAGATGGCTACCTCTCTTCACTGTCAGCCAG 638
Db 301 CTTGCTACAGTTTGGATGGCTCTGTGAGATGGCTACCTCTCTTCACTGTCAGCCAG 360
QY 639 CCACAGTCTGGGGCACTGCTCCAGTGGCATGTTCCCGTACCCCGGCGAGCAGCAG 698
Db 361 CCACAGTCTGGGGCACTGCTCCAGTGGCATGTTCCCGTACCCCGGCGAGCAGCAG 420
QY 699 GGGCAACCCCGCTTCTCATGCGCAGCCCTAGCCGCTCATGAGCCACACCCGGTCA 758
Db 421 GGGCAACCCCGCTTCTCATGCGCAGCCCTAGCCGCTCATGAGCCACACCCGGTCA 480
QY 759 ACCAAGGAGCTTACACAGCCTCTGTGCACTCTGTGCACTCTGTGCACTCTGTGCA 818
Db 481 ACCAAGGAGCTTACACAGCCTCTGTGCACTCTGTGCACTCTGTGCACTCTGTGCA 540
QY 819 GGCTACCCCTTACCCACAGAGTACGGGCACTCTTACCAAGAGCTCCGTTCTACCACTTC 878
Db 541 GGCTACCCCTTACCCACAGAGTACGGGCACTCTTACCAAGAGCTCCGTTCTACCACTTC 600
QY 879 TCCTCCACCCAGCCGGGCTGTGTGCTCCGGCAAGCAAGTGTACCTGTGCAAGCGCC 938
Db 601 TCCTCCACCCAGCCGGGCTGTGTGCTCCGGCAAGCAAGTGTACCTGTGCAAGCGCC 660
QY 939 CTTTGGCTGAATTTTCAACCGGCAACCAAGAGATGATATCAACCAACAGGGAAGGCG 998
Db 661 CTTTGGCTGAATTTTCAACCGGCAACCAAGAGATGATATCAACCAACAGGGAAGGCG 720
QY 999 ATGTTTCTTTTAAATTTTAAATTTCTGGTCTCGATCCGAGGCTCATTAACAATTT 1058
Db 721 ATGTTTCTTTTAAATTTTAAATTTCTGGTCTCGATCCGAGGCTCATTAACAATTT 780
QY 1059 TTTTGGATGTATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGCAAGTGGTT 1118
Db 781 TTTTGGATGTATTTTGGCGGATCCCAATCACTGAGGTTTCAAGAGCAAGTGGTT 840
QY 1119 CTTTGGCGGAAGCGGACCAATGTGCAAG---GAAATGGGTCTAATGATCCGGA 1174
Db 841 CTTTGGCGGAAGCGGACCAATGTGCAAGGCAAGGCAAGGATCGGTCTAATGATCCGGA 900
QY 1175 TTTCCGCCAACACTGGGGCTCACTGATGCGCCAAAGAAATCTCTTTTGGAAATTTAAACT 1234
Db 901 TTTCCGCCAACACTGGGGCTCACTGATGCGCCAAAGAAATCTCTTTTGGAAATTTAAACT 960
QY 1235 TACGAACAACAAGAGCTTCAAAATTAACAATGAGGAGATGTGTGTTTACAGTCTTGCA 1294
Db 961 TACGAACAACAAGAGCTTCAAAATTAACAATGAGGAGATGTGTGTTTACAGTCTTGCA 1020
QY 1295 CAAGTACCAAGCCCGCTGCGATGTGTGGAATGAAAGAGAGCGCAAGGAGAGACATAG 1354
Db 1021 CAAGTACCAAGCCCGCTGCGATGTGTGGAATGAAAGAGAGCGCAAGGAGAGACATAG 1080
QY 1355 CCAGCCCGGCGCGTGCAGACGTTCACTTCCCTGAGACTCAGTTCAATGCGCGTCAACGC 1414
Db 1081 CCAGCCCGGCGCGTGCAGACGTTCACTTCCCTGAGACTCAGTTCAATGCGCGTCAACGC 1140
QY 1415 CTACCAGAACACGAGATTTTACAACAATAAATAGATCAACACCTTTTGGAAAAAGATT 1474
Db 1141 CTACCAGAACACGAGATTTTACAACAATAAATAGATCAACACCTTTTGGAAAAAGATT 1200
QY 1475 TGGGATTAATTTATGACAGATCTACACCGGCTGTGACATGAGACCGCTTACCCCTCGGC 1534
Db 1201 TGGGATTAATTTATGACAGATCTACACCGGCTGTGACATGAGACCGCTTACCCCTCGGC 1260
QY 1535 CAACGACTGCGCGCTGCAAGTGTGTGCGCGGCGCGCTAGCGCATGCGCGGCTTT 1594
Db 1261 CAACGACTGCGCGCTGCAAGTGTGTGCGCGGCGCGCTAGCGCATGCGCGGCTTT 1320
QY 1595 CCTGACGAGCAAGTTCTGTAGAGCACTTACCGCAAGGCGCGCTTCAACCGCGGCGCGCGCG 1654
Db 1321 CCTGACGAGCAAGTTCTGTAGAGCACTTACCGCAAGGCGCGCTTCAACCGCGGCGCGCGCG 1380

QY 1595 GGGCCCCGGGCGGGTACCGACCGAGAGTGCAGCACCAACGAGGCTGTGCGCGCA 1714
Db 1381 GGGCCCCGGGCGGGTACCGACCGAGAGTGCAGCACCAACGAGGCTGTGCGCGCA 1440
QY 1715 GCAAGCGAGAGCAACCGGCGCGGCTTCCGCGCAACGCTGTGTTTGTGACCGCGCGCAACAA 1774
Db 1441 GCAAGCGAGAGCAACCGGCGCGGCTTCCGCGCAACGCTGTGTTTGTGACCGCGCGCAACAA 1500
QY 1775 CCGGCTGAGACTTGGCGGCTTGTGGCTTATGACACGGCCACCGACTTTCGGGCGCAACGGCGC 1834
Db 1501 CCGGCTGAGACTTGGCGGCTTGTGGCTTATGACACGGCCACCGACTTTCGGGCGCAACGGCGC 1560
QY 1835 CAAGCTGCTCTTACGCGGCGGCGGAGTAAAGGAGCTGCGCGTGCAGGCTGCAGGCTG 1894
Db 1561 CAAGCTGCTCTTACGCGGCGGCGGAGTAAAGGAGCTGCGCGTGCAGGCTGCAGGCTG 1620
QY 1895 CACTGGCGCGCGCTGCGCTTACTACGCGCAACCGTGTGGGCTGGGGCGCGCACTCCCC 1954
Db 1621 CACTGGCGCGCGCTGCGCTTACTACGCGCAACCGTGTGGGCTGGGGCGCGCACTCCCC 1680
QY 1955 GCAATATGCGGCAACCAAGTGGGCTGGTGTGCTGCTGTGCGCCAAACGGCGCGCGC 2014
Db 1681 GCAATATGCGGCAACCAAGTGGGCTGGTGTGCTGCTGTGCGCCAAACGGCGCGCGC 1740
QY 2015 CGCGCGCGCATGGCGCGGCGCAATCCCTTACTGAGGCGAGGAGCGCAAGGCGCTGCGCGC 2074
Db 1741 CGCGCGCGCATGGCGCGGCGCAATCCCTTACTGAGGCGAGGAGCGCAAGGCGCTGCGCGC 1800
QY 2075 CGAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCAAGCGCAAGCGCTGCGCA 2134
Db 1801 CGAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCAAGCGCAAGCGCTGCGCA 1860
QY 2135 TTTCAAGCTGATTCAGAGAGCTTCTCTGATTAAGTCCATGCACTCCAGCCACTGGGGAT 2194
Db 1861 TTTCAAGCTGATTCAGAGAGCTTCTCTGATTAAGTCCATGCACTCCAGCCACTGGGGAT 1920
QY 2195 TTTACGAGCAGGCGCAAGCGGAGATCTGCGCGCGCGCAACGCGCGCTGTGCGAGATT 2254
Db 1921 TTTACGAGCAGGCGCAAGCGGAGATCTGCGCGCGCGCAACGCGCGCTGTGCGAGATT 1980
QY 2255 GTCCCGCTCAAGACGAGGTGTGCGCGCGCGCGCGCGCGCGCGCAAGAACTGCGCAAGGA 2314
Db 1981 GTCCCGCTCAAGACGAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCAAGAACTGCGCAAGGA 2040
QY 2315 CATTAGCGGCTAATATGCTTCTACTGCAACGTAAGGCGCGCTTCCGCGCGCGCGCGCGC 2374
Db 2041 CATTAGCGGCTAATATGCTTCTACTGCAACGTAAGGCGCGCTTCCGCGCGCGCGCGCGC 2100
QY 2375 GCGCGGCGCGGAGCA 2434
Db 2101 GCGCGGCGCGGAGCA 2160
QY 2435 CTCTCTCTTGGCAACCACTGATTTTATGACCTCGATGGCGGTGCGAGCAATAG 2494
Db 2161 CTCTCTCTTGGCAACCACTGATTTTATGACCTCGATGGCGGTGCGAGCAATAG 2220
QY 2495 TGCAGGTCTCGAGCGGATTTTAACTTTTTCAGAGAGTGTGCAATTAGCTCAC 2554
Db 2221 TGCAGGTCTCGAGCGGATTTTAACTTTTTCAGAGAGTGTGCAATTAGCTCAC 2280
QY 2555 CGACCTTCAACTTTGCTGTAACCTTTTGTGTTTCTACTACTCTTCTGTGAGATT 2614
Db 2281 CGACCTTCAACTTTGCTGTAACCTTTTGTGTTTCTACTACTCTTCTGTGAGATT 2340
QY 2615 ATCTCTCTCAATTCCTCTCCCGTGTCTTCTTACTCTCACTCTTCTTCTTCTGTA 2674
Db 2341 ATCTCTCTCAATTCCTCTCCCGTGTCTTCTTACTCTCACTCTTCTTCTTCTGTA 2400
QY 2675 ATGAAACTTTCACCTTTAGAGACCTGGGAGTCTGTGAGGAGAGAGCAATTCGAGC 2734
Db 2401 ATGAAACTTTCACCTTTAGAGACCTGGGAGTCTGTGAGGAGAGAGCAATTCGAGC 2460

[illegible]

```

RESULT 2
US-09-949-016-16576
; Sequence 16576, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16576
; LENGTH: 12681
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16576

Query Match      53.2%; Score 1695.4; DB 3; Length 12681;
Query Local Similarity 99.9%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Dn      Qy      Db
1490 CACGATCTACACCGCGCTGTGACATGAGCCGCGCTGACCCCTCGCGCCCAACGACTGCGCGG 1549
    |||
8985 CAGGATCTACACCGCGCTGTGACATGAGCCGCGCTGACCCCTCGCGCCCAACGACTGCGCGG 9044
1550 CTGCGCATGCTGCGCGCGCGCGCGCTACGCGCATGCGCGCGCGCTCTTTCTGTGACGACAGTT 1609
9045 CTGCGCATGCTGCGCGCGCGCGCGCTACGCGCATGCGCGCGCGCTCTTTCTGTGACGACAGTT 9104
1610 CGGAGCACTACCGCCCAAGCGCGCTTCAACCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGG 1669
9105 CGTAGCACTACCGCCCAAGCGCGCTTCAACCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGG 9164

```

QY	1670	TATGACACCGCAGCTGTCGCGACACCAAGGGGTGTCGTGCGCGAGACAGGGCCGAGACACC	1729
Db	9165	TATGAGACCCGACGTCGCGACACCAAGGGGTGTCGTGCGCGAGACAGGGCCGAGACACC	9224
QY	1730	GGCGGCGCCCTCGCGCAACGCTGGTTTGTAACGCGCGGCAACCAACCGGCTGACTTCGC	1789
Db	9225	GGCGGCGCCCTCGCGCAACGCTGGTTTGTAACGCGCGGCAACCAACCGGCTGACTTCGC	9284
QY	1790	GAGCTCGGCTATGACACGCGCAACGACTTCGCGGGCAACGCGGCAAGCTGCTCTTAA	1849
Db	9285	GAGCTCGGCTATGACACGCGCAACGACTTCGCGGGCAACGCGGCAAGCTGCTCTTAA	9344
QY	1850	CGCGGCGGCGGCGGTGAAGCGCTCCCGCTGAGGCTGCAAGCTGACTGGCGCGGCT	1909
Db	9345	CGCGGCGGCGGCGGTGAAGCGCTCCCGCTGAGGCTGCAAGCTGACTGGCGCGGCT	9404
QY	1910	CGGCTACTAGCGGACCCGCTCGGCTGCGGGCGCGCGAGTCCCGCGAGTACTCGGAC	1969
Db	9405	CGGCTACTAGCGGACCCGCTCGGCTGCGGGCGCGCGAGTCCCGCGAGTACTCGGAC	9464
QY	1970	CAAGTCGGGCTCGGTCGTCGCTGCGCCAAACGCGCGCGCGCGCGCATGCG	2029
Db	9465	CAAGTCGGGCTCGGTCGTCGCTGCGCGCCAAACGCGCGCGCGCGCATGCG	9524
QY	2030	CGGCGCAATCTCCTACCTGGGCGAGAGGCGAGGCGCTGGCGCGCGGAGCGCTCGCGCT	2089
Db	9525	CGGCGCAATCTCCTACCTGGGCGAGAGGCGAGGCGCTGGCGCGCGGAGCGCTCGCGCT	9584
QY	2090	GCGGCGCGGCGCGCGGAGAGCGCAAGCGCAAGGACCTGTCGAAATTCAGCTGGATGA	2149
Db	9585	GCGGCGCGGCGCGCGGAGAGCGCAAGCGCAAGGACCTGTCGAAATTCAGCTGGATGA	9644
QY	2150	GACGCGCTCTCGATCAAGTCATCGACTTCAGCGACTCGGAGATTYACGAGCAGGCCAA	2209
Db	9645	GACGCGCTCTCGATCAAGTCATCGACTTCAGCGACTCGGAGATTYACGAGCAGGCCAA	9704
QY	2210	GCGGAGGCGGATCTCGCGGCGCGACACGCGCGTGTCCAGAGTTGTGTCGCGCTCAAG	2269
Db	9705	GCGGAGGCGGATCTCGCGGCGCGACACGCGCGTGTCCAGAGTTGTGTCGCGCTCAAG	9764
QY	2270	CGAGGTCGTGCGCAGCGGAGCTGCGAGAAATCGCGCAGAGCAATTAGCGGCTACTA	2329
Db	9765	CGAGGTCGTGCGCAGCGGAGCTGCGAGAAATCGCGCAGAGCAATTAGCGGCTACTA	9824
QY	2330	TGGCTTCTACCTGCGACAGTAGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCG	2389
Db	9825	TGGCTTCTACCTGCGACAGTAGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCG	9884
QY	2390	CCGAGCGAGCGCCCTCAGAGCTCTTCCCGAGTCCGCTCCCGACACTCTCTTGGCGAC	2449
Db	9885	CCGAGCGAGCGCCCTCAGAGCTCTTCCCGAGTCCGCTCCCGACACTCTCTTGGCGAC	9944
QY	2450	CGACTCATTTTATTTGACCTTCGATGGCGGTCTGCGAGCGAATAAGGAGGTCTCGAGC	2509
Db	9945	CGACTCATTTTATTTGACCTTCGATGGCGGTCTGCGAGCGAATAAGGAGGTCTCGAGC	10004
QY	2510	GATGATTTTAACTTTTGGACAGAGCTCTGCAATTAAGTCTACCGACTTCAACTTGG	2569
Db	10005	GATGATTTTAACTTTTGGACAGAGCTCTGCAATTAAGTCTACCGACTTCAACTTGG	10064
QY	2570	CTGTAAACCTTTGGTTTCTCTACTTACTCTCTCTTCTGTGGAGTTATCTCTCAATTG	2629
Db	10065	CTGTAAACCTTTGGTTTCTCTACTTACTCTCTCTTCTGTGGAGTTATCTCTCAATTG	10124
QY	2630	CCCTTCCCTCTGTCTTTCTCTTACTCTCTCTCTCTCTTCTGTAATGAACCTTCAAC	2689
Db	10125	CCCTTCCCTCTGTCTTTCTCTTACTCTCTCTCTCTCTTCTTGTGAATGAACCTTCAAC	10184
QY	2690	TTTATGAGACCTGGGAGAGTCTGTGACAGGACAGCATTCGACCCGCAAGTCTCGGC	2749
Db	10185	TTTATGAGACCTGGGAGAGTCTGTGACAGGACAGCATTCGACCCGCAAGTCTCGGC	10244

[illegible]

```

RESULT 3
US-09-775-398-37/C
; Sequence 37, Application US/09775398
; Patent No. 6893820
; GENERAL INFORMATION:
; APPLICANT: Plas, Christoph
; TITLE OF INVENTION: Detection of Methylated CpG Rich Sequences Diagnostic for Maligna
; FILE REFERENCE: 22727/04075
; CURRENT APPLICATION NUMBER: US/09/775,398
; CURRENT FILING DATE: 2001-01-04
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens 3.B.55
US-09-775-398-37

Query Match          16.1%; Score 514; DB 3; Length 514;
Beet Local Similarity 100.0%; Pred. No. 6.5e-101;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1504  GCTGTGAATGAGCCGCTGACCCCTCGCCCAAGATTCGCGCGCTGCAATGTCG 1503
Db       514  GCTGTGAATGAGCCGCTGACCCCTCGCCCAAGATTCGCGCGCTGCAATGTCG 455
QY      1504  CCGGGGCCCGCTAAGCCCAATGCGCGGCTCTTTCCTGACAGACCAAGTTCGTAGCAACTAAG 1623
Db       454  CCGGGGCCCGCTAAGCCCAATGCGCGGCTCTTTCCTGACAGACCAAGTTCGTAGCAACTAAG 395
QY      1624  CCAAGGCCCGCTTTCACCCCGGGGCGGGGCGCGGGCCCGGGCCGGGTACGACCGCAGCG 1683
Db       394  CCAAGGCCCGCTTTCACCCCGGGGCGGGGCGGGGCGCGGGCCCGGGGTACGACCGCAGCG 335
QY      1684  TGCCGCACACCAACGAGGCTGTGTGTGCGCGACAGCGACGAGACCCCGGGCGCGCTTGC 1743
Db       334  TGCCGCACACCAACGAGGCTGTGTGTGCGCGACAGCGAGGACCCCGGGCGCGCTTGC 275
QY      1744  CGCAACGCTGTTTGTGACGCGCGCAACACCGGCTGACTTCGCGGCTTCGCGCTATG 1803

```

Db	214	CGCAACGCTGGTGTGTGAACCGCGGCAACACCAGCTGACTTCGGCGCTCTAAG	215
QY	1804	ACACGGCCACGAGACTTTCGCGGCGCAACGCGGCCACGCTGCTCTTAAACGCGCGCGCGGCG	186
Db	214	ACACGGCCACGAGACTTTCGCGGCGCAACGCGGCCACCGCTGCTCTTAAACGCGCGCGCGGCG	155
QY	1884	TGAAGGCGCTGCGCGCTGCAAGCTGCAAGGCTGCACTGGCGCGCGCGCTCACTACGCGG	192
Db	154	TGAAGGCGCTGCGCGCTGCAAGCTGCAAGGCTGCACTGGCGCGCGCGCTCACTACGCGG	95
QY	1924	ACCCGTCGGGCTGGGCGCGCCGCACTGCCCGCACTACGCGGCAACAAGTCGGGCTCGG	198
Db	94	ACCCGTCGGGCTGGGCGCGCCGCACTGCCCGCACTACGCGGCAACAAGTCGGGCTCGG	35
QY	1984	TGCTGCTCTGCTGCTGCGCCCAACAGCGCGCGGCGCGC	2017
Db	34	TGCTGCTCTGCTGCTGCGCCCAACAGCGCGCGCGCGC	1

```

/ RESULT 4
/ US-09-513-999C--885
/ Sequence 885, Application US/0951399C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclert, A.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ Patent No. 6783961
/ FILE REFERENCE: 59..US2.REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ CURRENT FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/122,487
/ PRIOR FILING DATE: 1998-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
/ SEQ ID NO 885
/ LENGTH: 388
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 117..386
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4
/ OTHER INFORMATION: w=a or t
/ US-09-513-999C--885

```

Query Match	12.1%;	Score 384.6;	DB 3;	Length 388;
Best Local Similarity	99.7%;	Pred. No. 3.9e-73;		
Matches 384;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0
QY	190	TTTGTGGTTGAAGCTTTCTGTCTAGAGAGGGGTCTGTGATTTCTAGTTTATGATA	249	
Db	4	WTTGCTGTTGAAGTCTTTCTGTCTAGAGAGGGGTCTGTGATTTCTAGTTTATGATA	63	
QY	250	AATAGACTTTAAAAACGAGGACGGGAGGGCGAGTGTTCAGGTTCTAGACTATGACG	309	
Db	64	AATAGACTTTAAAAACGAGGACGGGAGGGCGAGTGTTCAGGTTCTAGAGCTATGACG	123	
QY	310	TGGAGACATGCTTTCTCCTTTATCAGCTCTCCAGAAATTTCTCAATGTGAGCAGA	369	
Db	124	TGGAGACATGCTTTCTCCTTTATCAGCTCTCCAGAAATTTCTCAATGTGAGCAGA	183	
QY	370	GCTACCCACATTCAGGCGGATCCGAGCTTGTCTGACGATCATCCATTATTCGACCA	429	
Db	184	GCTACCCACATTCAGGCGGATCCGAGCTTGTCTGACGATCATCCATTATTCGACCA	243	
QY	430	CTGAAACCTGGAGAGAGTTCACCTTTGAAAAAATTAACGAGGGGATGAGAACTCAGT	489	
Db	244	CTGAAACCTGGAGAGAGTTCACCTTTGAAAAAATTAACGAGGGGATGAGAACTCAGT	303	

QY 490 CAGATACAGACAAATTTCTGACTCCAAAGACTCACCAGGGAGCTCCAGAGAGTAAAC 549
 DB 304 CAGATACAGACAAATTTCTGACTCCAAAGACTCACCAGGGAGCTCCAGAGAGTAAAC 363
 QY 550 TCTCTCTGCTCTGAGAGGGGGCTC 574
 DB 364 TCTCTCTGCTCTGAGAGGGGGCTC 388

RESULT 5
 US-09-188-811-1
 ; Sequence 1, Application US/09188811
 ; Patent No. 6037148
 ; GENERAL INFORMATION:
 ; APPLICANT: Rhododoust, Mehran
 ; TITLE OF INVENTION: NOVEL MTX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: MNI-046CP
 ; CURRENT APPLICATION NUMBER: US/09/188,811
 ; CURRENT FILING DATE: 1998-11-09
 ; EARLIER APPLICATION NUMBER: 09/163,116
 ; EARLIER FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1704
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1701)
 US-09-188-811-1

Query Match 11.8%; Score 374.6; DB 3; Length 1704;
 Best Local Similarity 68.8%; Pred. No. 9.1e-71;
 Matches 515; Conservative 0; Mismatches 234; Indels 0; Gaps 0;
 QY 912 GCACAGGTGTACCTGTGCAACAGGCCCTTTGGCTGAAATTTACCGGACCAAAAGGAG 971
 DB 385 GCCCAGCTTACTCTGTGCAACCGGCTCTGTGGCTCAATTTACCGGACCAAAAGTGA 444
 QY 972 ATGATCATCAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1031
 DB 445 ATGATCATTAACGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504
 QY 1032 CTCGATCCCAAGGCTATTAACAATTTTGGAGATGATTTGGCGGATCCCAATCAG 1091
 DB 505 CTCGATCCCAAGGCTATTAACAATTTTGGAGATGATTTGGCGGATCCCAATCAG 564
 QY 1092 TGGAGGTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1151
 DB 565 TGGCGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 624
 QY 1152 AATCGGCTATATGATCGGATTCCTCCCAACACTGGGAGCTGATGAGGAGGAGGAG 1211
 DB 625 AACAAATGATGATTCACCAAGAGTCTCTAATATGATGATGATGATGATGATGATGAT 684
 QY 1212 ATCTCTTTGGAATAAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAA 1271
 DB 685 ATTTCAATGGGAAATTAATAAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTA 744
 QY 1272 ATGAGGTTTCAAGTCTTGAACAAGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1331
 DB 745 ATGATGATCTTAATCAATCTTAACCAATTAACCAATTAACCAATTAACCAATTAAC 804
 QY 1332 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1391
 DB 805 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 864
 QY 1392 ACTCAGTCAATCGGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451
 DB 865 AGCAATTCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 924

QY 1452 CACAACCTTTTGCAAAAAGATTTGGGATTAATATGACAGATCTACACCGGCTGTGAC 1511
 DB 925 CATTAACCTTTTGCAAAAAGGCTTCAGAGACATATATATTCATGTACACGGCTTGAGA 984
 QY 1512 ATGAGACGCTGACCCCTGCGCCCAACGACTCGCGGCTGCGAGATGCGCGGGGCC 1571
 DB 985 AATGACAGGTTAACTCATCTCCACAGATTTCTGATTCCTCAATGATTTCTTGA 1044
 QY 1572 CGTACGCCATGCGGCTCTTTCTGAGAGACCAAGTTCGTGAGCACTACGCCAAGGCC 1631
 DB 1045 GATCGTACCGCGCTTCAATCTTCTTCCGAGGCTTTTGCAACACTTAACCTTAAGCC 1104
 QY 1632 CGCTTCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1660
 DB 1105 CGTATTAATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1133

RESULT 6
 US-09-189-760-1
 ; Sequence 1, Application US/09189760
 ; Patent No. 6031078
 ; GENERAL INFORMATION:
 ; APPLICANT: Rhododoust, Mehran
 ; TITLE OF INVENTION: NOVEL MTX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: MNI-046CP2
 ; CURRENT APPLICATION NUMBER: US/09/189,760
 ; CURRENT FILING DATE: 1998-11-10
 ; EARLIER APPLICATION NUMBER: 09/163,116
 ; EARLIER FILING DATE: 1998-09-29
 ; EARLIER APPLICATION NUMBER: 60/089,467
 ; EARLIER FILING DATE: 1998-06-16
 ; EARLIER APPLICATION NUMBER: (PENDING)
 ; EARLIER FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2494
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (164)..(1714)
 US-09-189-760-1

Query Match 11.8%; Score 374.6; DB 3; Length 2494;
 Best Local Similarity 68.8%; Pred. No. 1e-70;
 Matches 515; Conservative 0; Mismatches 234; Indels 0; Gaps 0;
 QY 912 GCACAGGTGTACCTGTGCAACAGGCCCTTTGGCTGAAATTTACCGGACCAAAAGGAG 971
 DB 385 GCCCAGCTTACTCTGTGCAACCGGCTCTGTGGCTCAATTTACCGGACCAAAAGTGA 444
 QY 972 ATGATCATCAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1031
 DB 445 ATGATCATTAACGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504
 QY 1032 CTCGATCCCAAGGCTATTAACAATTTTGGAGATGATTTGGCGGATCCCAATCAG 1091
 DB 505 CTCGATCCCAAGGCTATTAACAATTTTGGAGATGATTTGGCGGATCCCAATCAG 564
 QY 1092 TGGAGGTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1151
 DB 565 TGGCGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 624
 QY 1152 AATCGGCTATATGATCGGATTCCTCCCAACACTGGGAGCTGATGAGGAGGAGGAG 1211
 DB 625 AACAAATGATGATTCACCAAGAGTCTCTAATATGATGATGATGATGATGATGATGAT 684
 QY 1212 ATCTCTTTGGAATAAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAA 1271
 DB 685 ATTTCAATGGGAAATTAATAAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTA 744
 QY 1272 ATGAGGTTTCAAGTCTTGAACAAGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1331
 DB 745 ATGATGATCTTAATCAATCTTAACCAATTAACCAATTAACCAATTAACCAATTAAC 804
 QY 1332 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1391
 DB 805 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 864
 QY 1392 ACTCAGTCAATCGGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451
 DB 865 AGCAATTCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 924

Oy	1212	ATGCGGTTTAAAGTCTCTTGACAAAGTACAGAGCCCGCTGCAATGTGTGAAGTGAAC	1331
Db	758	ATGATAGTCTTAACATCCCTACCAAAATACCAACCCCACTGCATATGTGGAACTGA	817
Oy	1332	GAGCAGGCAAGAGAGACACTAGCCAGGCGGGCGCGTGACAGAGTTCACTTCCCTAG	1391
Db	818	GAGGATGGCGTGGAGGACTTGAATAGCCCTAAGAACCCAGACTTTTAACTTCTCAGAA	877
Oy	1392	ACTCAGTTCAATGCCGTACCGGCTACAGAAACAGGATATTATACCAACTGAAATAGAT	1451
Db	878	AACGCAATTCATGTGACGTACTGCTTACCAAAACACGATATTACTCAACTAAAGTTGAT	937
Oy	1452	CACAACTCTTTTGCAAAAGATTTGGGATATTATGACAGATCTAACCGGCTGTAC	1511
Db	938	CATAACCCCTTTGCAAAAGGCTTGAGAAACAATATGATTCATGTACATCAACCGCTTCAAA	997
Oy	1512	ATGACCGGCTGAGCCCTCGGCCCAAGACACTGGCGGCGCTGCAGATCGTGCCCGGGGCC	1571
Db	998	AATGACAGGTTAATCTTCATCTTCCACGAGATTTCTTAATCCATCAAGATTGTCCCTGGA	1055
Oy	1572	CGCTACGCAATGCGCGGCTCTTTCTTGACAGACCAATTGAGAGCACTAACGCCAAGGCC	1631
Db	1058	GATCGGTAACGGGTTCAATCCTCTTTCCCGAGAGCCCTTTGTCAACACTTTAAGCTAAGCC	1117
Oy	1632	CGCTTCCACCCGGGCGCGGGCGCGGGGCC	1660
Db	1118	CGCTATTATTAATGCGAGAGAACTGTGCC	1146

```

RESULT 7
US-09-514-422-1
; Sequence 1, Application US/09514422
; Patent No. 6291193
; GENERAL INFORMATION:
APPLICANT: Khodadoust, Mehran
TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: MN1-046CP2
CURRENT FILING DATE: US/09/514,422
CURRENT FILING DATE: 2000-02-28
PRIORITY APPLICATION NUMBER: US/09/189,760
PRIORITY FILING DATE: 1998-11-10
PRIORITY APPLICATION NUMBER: 09/163,116
PRIORITY FILING DATE: 1998-09-29
PRIORITY APPLICATION NUMBER: 60/089,467
PRIORITY FILING DATE: 1998-06-16
PRIORITY APPLICATION NUMBER: (PENDING)
PRIORITY FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 2494
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(1714)
US-09-514-422-1

Query Match      11.8%; Score 374.6; DB 3; Length 2494;
Best Local Similarity 68.8%; Pred. No. 1e-70;
Matches 515; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY      912 GCACAGGCTACTTGTGCACAAGCGCCCTTTGGCTGAATTTCAACCGGCAACCAAGGAG 971
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      398 GCCACGCTCTACCTCTGTGCACCGCGCTCTGTGGCTCAATTCCACCGGCCAACCAACTGAG 457
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      972 ATGATCATACCACCAACAGGGAAGGCGCAGTGTTCTTTTTTAAGTTTAAATTTCTGGT 1031
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      458 ATGATCATATTACGAACACAGGCGAGGCGCATGTTCTCTTCCTTAGGCTTCAACATAAAGCA 517
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      1032 CTCGATCCCAGGCTCATTAACATATTTTGTGATGTGATTTTGGCGGATCCCACATCAC 1091
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db	518	CTCAATCCCACTGCCCACTACATGTGTTCTGTAGAGGTGTGTGTGGCGAAGCCCAACAC	577
Qy	1092	TGAGAGTTTCAAGAGAGCAATGGTTCCTTCGCGCAAGCGAGACCAATGTGCAAGA	1151
Db	578	TGGCGCTTCCAGGGGGGCAATGGGTGACCTGTGGCAAGCGACATATGACGAGGC	637
Qy	1132	AATGGGTCTAATGATCCGGATTCCCAACACTGGGGGTCTACTGTATGGCCCAAGAA	1211
Db	638	AACAAAATGTATGTTCCACCGAAGTCTCTAATACTGTTTCCACTGTGATGAAACAGGAG	697
Qy	1212	ATCTCTTTTGGAAATTTAAATCTACGAAACAAGAGCTTCAATATACATATGGCAG	1271
Db	698	ATTTATTCGGGAATTTAAATCTACCAATATACAAAGGCGCAATATACAAACACCCAG	757
Qy	1272	ATGTGTGTTTACAGTCTTGTGACAAATACAGCCCCGCGCTGCATGTGTGAAGTGAAC	1331
Db	758	ATGATAGTCTTCAATCTCTTACACAAATCCAAACCCCGACTGATTTGTGAAGTTACA	817
Qy	1332	GAGGACGGACGAGAGACCTAGCCAGCCCGGCGGTGACAGCGTTCACTTTCCCTGAG	1391
Db	818	GAGGATGGGTGAGAGACTTGAATGAGCCCTCAAGACCAGCTTTTACTCTTAGAA	877
Qy	1392	ACTCAGTTCATGGCGGTCAACCGCTACACAGAACACAGGATATTACAACTGAAATAGAT	1451
Db	878	ACGCATTCATTGCAGTGACTGTGCTTACCAAAACCGGATTTACTCAACTAAAGATTGAT	937
Qy	1452	CACAACTCTTTTGGCAAAAGATTTGGGAAATTAATGACAGATCTTACACCGGCTGTAC	1511
Db	938	CATAACCCCTTTGGCAAAAGGCTTCAGAGACCACTATGATTTCCATGTACACCGCTTAGAA	997
Qy	1512	ATGGAACGGCTGACCCCTCGGCCAACGACTGGCGCGCTCCGCAAGATCGTGCCCGGGGCC	1571
Db	998	AATGACAGTTTATCTTCATCTCCACGGAATTCCTAGATCCCATCAGATATGTCTTGGA	1057
Qy	1572	CGCTACGCGCATGCGCGGCTCTTCTCTGACAGACAGATTGCTGAGCACTACGCCAAGGC	1631
Db	1058	GGTCGGTACGGGTTTCAATCTTCTTCCCGAGAGCCCTTTGTCAACCTTACTCAAGCC	1117
Qy	1632	CGCTTCCACCCGGGCGCGGGCGCGGCC	1660
Db	1118	CGCTATTATATGGCGAGAGAACCGTGCC	1146

```

      RESULT 8
US-09-949-016-3434
; Sequence 3434, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 3434
; LENGTH: 2575
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3434

Query Match          9.9%; Score 316.2; DB 3; length 2575;
Best Local Similarity 61.2%; Pred. No. 3,8e-58;
Matches 510; Conservative 0; Mismatches 323; Indels 0; Gaps 0;

```



```

Db      404  CCGGCGCCGCGAGCCGCTTCTTGAAGCTACGCGTACCGCGCCGCGAGCCCGCGGCG 463
Qy      756  ATCAACCAAGAGGAGCTTCAACAGAGCTTCTTCACTCTTCACTCTTCACTCTTCACT 815
Db      464  GGGTTCCTCCCGGCGGCGGAGGCTTCTTCCGCGCGCGCGCGGAGCGGCGGCTTCAAC 523
Qy      816  GCGGCTACCTCTTCACTTCAAGAGCTTCTTCACTCTTCACTCTTCACTCTTCACT 875
Db      524  GCGGAGGAGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 583
Qy      876  TTCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 935
Db      584  TACGGGCTACCGCGGAGCTTGAAGTGTGGGAACTGAGGAGTGTGGGAACTGAGGAG 643
Qy      936  CCGCTTGGCTGAAATTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 995
Db      644  CTGTGTGTGTCAAGTTTATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 703
Qy      996  CGCATGTTTCTTTTAAAGTTTAACTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1055
Db      704  CGCATGTTTCTTTTAAAGTTTAACTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 763
Qy      1056  ATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1115
Db      764  ATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 823
Qy      1116  GTTCTTGGCGGCAAGCGGAGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1175
Db      824  GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 883
Qy      1176  TCCCCCACTGTGGGCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1235
Db      884  TCCCCCACTGTGGGCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 943
Qy      1236  AGCAACAACAAGAGGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1295
Db      944  ACAACAACAAGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1003
Qy      1296  AGTACACAGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1355
Db      1004  AGTACACAGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1063
Qy      1356  CAGCCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1415
Db      1064  AAGCGTTCACACGCGATATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1123
Qy      1416  TACCAACAACAAGAGGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1475
Db      1124  TACCAACAACAAGAGGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1183
Qy      1476  CCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1528
Db      1184  CCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1236

```

RESULT 9
US-09-949-016-976
Sequence 976, Application US/09949016

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTRI, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 976
LENGTH: 2589
TYPE: DNA
ORGANISM: Human
US-09-949-016-976

Query Match 9.9%; Score 316.2; DB 3; Length 2589;
Best Local Similarity 61.2%; Pred. No. 3.8e-58;
Matches 510; Conservative 0; Mismatches 323; Indels 0; Gaps 0;

```

Qy      696  CCGGAGCAGCCCGCTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 755
Db      404  CCGGCGCCGCGAGCCGCTTCTTGAAGCTACGCGTACCGCGCCGCGAGCCCGCGGCG 463
Qy      756  ATCAACCAAGAGGAGCTTCAACAGAGCTTCTTCACTCTTCACTCTTCACTCTTCACT 815
Db      464  GGGTTCCTCCCGGCGGCGGAGGCTTCTTCCGCGCGCGCGCGGAGCGGCGGCTTCAAC 523
Qy      816  GCGGCTACCTCTTCACTTCAAGAGCTTCTTCACTCTTCACTCTTCACTCTTCACT 875
Db      524  GCGGAGGAGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 583
Qy      876  TTCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 935
Db      584  TACGGGCTACCGCGGAGCTTGAAGTGTGGGAACTGAGGAGTGTGGGAACTGAGGAG 643
Qy      936  CCGCTTGGCTGAAATTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 995
Db      644  CTGTGTGTGTCAAGTTTATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 703
Qy      996  CGCATGTTTCTTTTAAAGTTTAACTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1055
Db      704  CGCATGTTTCTTTTAAAGTTTAACTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 763
Qy      1056  ATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1115
Db      764  ATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 823
Qy      1116  GTTCTTGGCGGCAAGCGGAGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1175
Db      824  GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 883
Qy      1176  TCCCCCACTGTGGGCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1235
Db      884  TCCCCCACTGTGGGCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 943
Qy      1236  AGCAACAACAAGAGGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1295
Db      944  ACAACAACAAGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1003
Qy      1296  AGTACACAGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1355
Db      1004  AGTACACAGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1063
Qy      1356  CAGCCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1415
Db      1064  AAGCGTTCACACGCGATATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1123
Qy      1416  TACCAACAACAAGAGGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1475
Db      1124  TACCAACAACAAGAGGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1183
Qy      1476  CCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1528
Db      1184  CCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1236

```

RESULT 10
US-09-513-999C-22243
Sequence 22243, Application US/09513999C
Patent No. 6783961

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent .pm
; SEQ ID NO 22243
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-22243
```

```
Query Match          7.4%; Score 235; DB 3; Length 235;
Best Local Similarity 100.0%; Pred. No. 4.7e-41;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2923 TTTTATTCGTCAAGGACAAACAGTTCACTTATTTTTCCTTCTTCT
Db      1 TTTTATTCGTCAAGGACAAACAGTTCACTTATTTTTCCTTCTTCT
Qy      2993 TTGCTTTTCTTCTCTCTCTATCTTCTTCTCTCTTTTAATTTTCTGTGGA 3042
Db      61 TTGCTTTTCTTCTCTCTCTCTATCTTCTTCTCTTTTAATTTTCTGTGGA 120
Qy      3043 TAAATTTCTAAGAGCTCTAGAAACATGAAATACTCAGTAGAGTGAGGTTCCACTTCT 3102
Db      121 TAAATTTCTAAGAGCTCTAGAAACATGAAATACTCAGTAGAGTGAGGTTCCACTTCT 180
Qy      3103 CCTCAATCCGTTCGTAAGAAATATTACTATGTGCCCTTAATGACACAAATAGCTA 3157
Db      181 CCTCAATCCGTTCGTAAGAAATATTACTATGTGCCCTTAATGACACAAATAGCTA 235
```

```
RESULT 11
US-09-949-016-656
; Sequence 656, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 656
; LENGTH: 1798
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-656
```

```
Query Match          4.7%; Score 150.6; DB 3; Length 1798;
Best Local Similarity 58.4%; Pred. No. 1.3e-22;
Matches 330; Conservative 0; Mismatches 214; Indels 21; Gaps 3;
```

```
Qy      919 TGTACTCTGCAACAGGCCCCCTTTGGCTGAATTTTCAACCGGACCAACGAGATGATCA 978
Db      302 TGAGCTTGAGAACCGGAGCTATGAGAGAGATTCAAGCTTCTGTGTGGAAACAGAAATGATCA 361
```

```
Qy      979 TCACCAAAAGAGGAGGCGCATGTTCTTTTAAAGTTTAAATTTCTGCTCGATC 1038
Db      362 TCACCAAAAGCTGGAGGCGCATGTTCTTCTGCTCCGAGTCTAGTCACTGCGCTTGAGAC 421
Qy      1039 CCAGGCTCATTAATATTTTGTGATGTGATTTTGGCGGATCCCAATCATCTGAGAGT 1098
Db      422 CCGAGGCCGCTACTTGTGTTCTTCTGATGTGATTCGGAGATGGGAGCTGCTACCGCT 481
Qy      1099 TTCAAGAGCAAAATGTGTTCTTGGCGCAAGGCGACCAATGTCAGAAATGGG 1158
Db      482 GGCAGGCGCGGCTGGAGGCCAGCGCAAGGAGAGCCCGCTGCTG--ACGTTG 538
Qy      1159 TCTATATGATCCGAGATTTCCCAACACTGGGCTCATGTAGTGGCCAGAAATCTTCT 1218
Db      539 TCTATATTCACCCGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
Qy      1219 TTGAAATTTAAATTTAATGCAACCAAGAGCTTCAATATACATATGGGAGATGCTG 1278
Db      599 TCCATCTGTCAACCTCACCAACAGACAGCTGAGACCCCAAGGCA-----CCTGA 649
Qy      1279 TTTTACAGTCTTGCACAAATGACAGCCCGCTGCATGTGTGAGATGAAAGAGAGC 1338
Db      650 TCTGCACTCATGACAAATGACCAACCCGCAATACCTAGTTCGGGACGCCAGCTCT 709
Qy      1339 GCACGAGGACACTAGCCAGCCCGCGCTGACAGAGTTCACTTCTCTGAGACTAGT 1398
Db      710 GCAGCGAG-----CACTGGGGGGGCAATGGCTCTCTTCCGCTTCCCGAGACCAAT 760
Qy      1399 TCATGCGGTCACCGCTTACCAAGACAGATATTATACAACTGAATATGATACAAAC 1458
Db      761 TCATCTCGTGAACAGCTTACCAAGACAGATATTATACAACTGAATATGACCAATC 820
Qy      1459 CTTTTCGAAAGATTTCCGGATTA 1483
Db      821 CTTTTCGAAAGGCTTCGGGAGAA 845
```

```
RESULT 12
US-09-949-016-2908
; Sequence 2908, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2908
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2908
```

```
Query Match          4.7%; Score 150.6; DB 3; Length 1799;
Best Local Similarity 58.4%; Pred. No. 1.3e-22;
Matches 330; Conservative 0; Mismatches 214; Indels 21; Gaps 3;
```

```
Qy      919 TGTACTCTGCAACAGGCCCCCTTTGGCTGAATTTTCAACCGGACCAACGAGATGATCA 978
Db      302 TGAGCTTGAGAACCGGAGCTATGAGAGAGATTCAAGCTTCTGTGTGGAAACAGAAATGATCA 361
Qy      979 TCACCAAAAGAGGAGGCGCATGTTCTTTTAAAGTTTAAATTTCTGCTCGATC 1038
Db      362 TCACCAAAAGCTGGAGGCGCATGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 421
```

1039 CCACGGCTCATTTACATATTTTGTGATGATTTTGGCGGATCCCATCTGAGAGT 1098
 422 CGAGAGCCCGCTACTTGTCTTCTTGATGATTCGCTGATGAGGCTGCTACCCCT 481
 1099 TTCAAGAGGCAAAATGGGTCTTGGCGCAAGGAGCAACAAATGCAAGGAATTCGGG 1158
 482 GGCAGAGGCGGCGCTGAGAGCCGACGCAAGAGAGCCCGCTGCTG---ACCGTG 538
 1159 TCTATGATCCGAGATTCCTCCCAACACTGAGGCTCATGTGATGCGCAAGAAATCTCTT 1218
 539 TCTATGATCCGAGATTCCTCCCAACACTGAGGCTCATGTGATGCGCAAGAAATCTCTT 598
 1219 TTGAAAATTTAAATCTTACGAAACAAAGAGCTTCAATATACATGAGGAGATGTTGG 1278
 599 TCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649
 1279 TTTTACAGTCTTGTGACAAATACCAAGCCCGCTGATGATGATGATGATGATGATGAT 1338
 650 TCTGCACTCCATGACAAATGATCAACCCGATACCTAGTTGGGAGCCGCTGCTT 709
 1339 GCACGAGGACACTAGCCGCGCGCGCTGCAAGCTTCACTTCCCTGAGACTCAAT 1398
 710 GAGCGAG-----CACTGGGGGGGATGAGCTTCTTCCGCTTCCCGAGACCAAT 1458
 1399 TCATGCGCTGACCGCTTACCAAGACCGATATTTACACACTGAAATATGATCAACAC 1458
 761 TCATGCGCTGACCGCTTACCAAGACCGATATTTACACACTGAAATATGATCAACAC 820
 1459 CTTTTCGAAAGAGATTTCCGGAATA 1483
 821 CCTTTCGAAAGAGATTTCCGGAATA 845
 RESULT 13
 US-09-949-016-4821
 / Sequence 4821, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / PRIOR FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / NUMBER OF SEQ ID NOS: 207012
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO: 4821
 / LENGTH: 2433
 / TYPE: DNA
 / ORGANISM: Human
 US-09-949-016-4821

Query Match 4.6%; Score 146; DB 3; Length 2433;
 Best Local Similarity 57.2%; Pred. No. 1.4e-21;
 Matches 335; Conservative 0; Mismatches 220; Indels 21; Gaps 3;

899 GGTGCGCGGCAAGACAGGTGATCTGTGCAACAGGCCCTTGGCTGAATTTTCACCG 958
 555 GGTGAGAGACGACCCCAAGGTGACCTGAGAGCTTAAGAACTTTGGATCAGTTTCACAA 614
 959 GCACCAAGAGAGATGATCATCAACCAAGAGGAGCGCATGTTCTTTTAAAGTT 1018
 615 GCGGGGACGAGATGATCATTAACCAAGTGGGAGGCGAATGTTCTTCATTTAAAGT 674
 1019 TAACATTTCTGCTGATGCCAGGCTCATTAATATTTTGTGATGATGATTTGGC 1078

675 GAGATGTTCTGGCTGATTAATAAAGCCAAATACATTTATTTGATGACATTAATAGCTGC 734
 1079 GATATCCCAATACATGAGATTTTCAAGAGCAATGGGTTCTTCCGCAAAAGCGACAC 1138
 735 TGAATGATGCTGATTAATTAATTTCAATTTCTCGGTGATGATGATGATGATGATGAT 794
 1139 CAATGTCGAAGAAATCGGCTCTATATGATCCGATTCCTCCCAACCTGGGCTGCTG 1198
 795 GGAATGTC---CAAGAGGATGATCATCAACCGGAGAGCCCGCTACTGGGAAACAGTG 851
 1199 GATGCGCCAGAAATCTCTTTTGAATAATTAATTAATTAATTAATTAATTAATTAAT 1258
 852 GATGTCGAAGATGCTGATTTTCAAAATCTGAAATCTCAACCAACCAATTTAG----- 904
 1259 TAACATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1318
 905 -ACAAATGATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 962
 1319 GGTGAGTGAAG 1378
 963 TGTAAAGCCCAATGATCATTTGAAAT-----CCCTATATGATCATTTCCGACATA 1013
 1379 CACTTTCCTGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438
 1014 CTTGTTCCCGAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073
 1439 ACTGAATATGATCAACCTTTTGAATAAGATTTGGGATAT 1484
 1074 GTTAAATATGATCAACCTTTTGAATAAGATTTGGGATAT 1119

RESULT 14
 US-09-949-016-4822
 / Sequence 4822, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / PRIOR FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / NUMBER OF SEQ ID NOS: 207012
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO: 4822
 / LENGTH: 2433
 / TYPE: DNA
 / ORGANISM: Human
 US-09-949-016-4822

Query Match 4.6%; Score 146; DB 3; Length 2433;
 Best Local Similarity 57.2%; Pred. No. 1.4e-21;
 Matches 335; Conservative 0; Mismatches 230; Indels 21; Gaps 3;

899 GGTGCGCGGCAAGACAGGTGATCTGTGCAACAGGCCCTTGGCTGAATTTTCACCG 958
 555 GGTGAGAGACGACCCCAAGGTGACCTGAGAGCTTAAGAACTTTGGATCAGTTTCACAA 614
 959 GCACCAAGAGAGATGATCATCAACCAAGAGGAGCGCATGTTCTTTTAAAGTT 1018
 615 GCGGGGACGAGATGATCATTAACCAAGTGGGAGGCGAATGTTCTTCATTTAAAGT 674
 1019 TAACATTTCTGCTGATGCCAGGCTCATTAATATTTTGTGATGATGATTTGGC 1078
 675 GATGATGTTCTGCTGATTAATAAAGCCAAATATTTTATGATGATGATTAATGATGCTG 734
 1079 GATTCCAATCACTGAGGTTTCAAGAGGCAAAATGGTTCCTTGGCGCAAGCGGACAC 1138

```

Db      735 TGATACGTGCTTATTAATTTACAAATTCCTCGGTGATGAGTGGCTGTGAAGCCGACCC 794
QY      1139 CAATGTGCAAGAAATCGGGTCTATATGATCCGGAATTCGCCAACACTGGGGCTCACTG 1198
Db      795 CGAAATGCG---CAAAAGAGATGTACATTCACCCGGACAGCCCGCTACTGGGGAAACAGTG 851
QY      1199 GATGGCGCAAGAAATCTCTTTGGAAATTTAACTTGAACCTTGAACCAACAAAGAGCTTCAA 1258
Db      852 GATGTCCAAAGTCGTACCTTTCCACAAACTGAACCTCACCAACCAATTTTCAG----- 904
QY      1259 TAACAATGGCGCAGATGTGTGTTTACAGTCTTGGACAAGTACCAAGCCCGCTGCATGT 1318
Db      905 --ACAAACATGATTTATCATATTTGAATCTCCATGCAAAATACAGAGCCCGGTTCCACAT 962
QY      1319 GGTGAAGTGAACGAGAGACGGCAGGAGACACTAGCCAGCCCGCGCGTGCAGAGCTT 1378
Db      963 TGTAGAGCGCAATGACATCTTGAACCT-----CCCTTATAGTACATTTGCGACATA 1013
QY      1379 CACTTCCCTGAGACTGATTCAGTCCGCTGACCGCTACCGAAGACAGATATTACACA 1438
Db      1014 CTGTTCCTCCGAACTGATTTATGCTGTGTGACTGATACAGAAATGATATAGATAACCA 1073
QY      1439 ACTGAAAATAGATCAAAACCTTTTGCAAAAGATTTGCGGATTAAT 1484
Db      1074 GTTAAATATGACAAACACCTTTTGCAAAAGTTTCGGGACACT 1119

```

```

RESULT 15
US-10-028-272-1
; Sequence 1, Application US/10028272
; Patent No. 6635481
; GENERAL INFORMATION:
; APPLICANT: Oligene Technologies
; TITLE OF INVENTION: TBX3 GENE AND METHODS OF USE
; FILE REFERENCE: 16U 104 R1
; CURRENT APPLICATION NUMBER: US/10/028,272
; CURRENT FILING DATE: 2001-12-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3113
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (524)..(2695)
; OTHER INFORMATION:
US-10-028-272-1

```

```

Query Match      4.6%; Score 146; DB 3; Length 3113;
Best Local Similarity 57.2%; Pred.No. 1.6e-21;
Matches 335; Conservative 0; Mismatches 230; Indels 21; Gaps 3;

```

```

QY      899 GGTGCCCGGCAAGCAGACGATGTAACCTGTGCAACAGGCCCTTTGGCTGAATTTACCG 958
Db      817 GGTGAGAGCGACGCCCAAGGTGACCTTGAGGCTAAAGAACTTTGGATCACTTTACAA 876
QY      959 GCACCAAAAGAGATGATCATCAACCAACAGGAAAGCGCATGTTCTTTTAAAGTTT 1018
Db      877 GCGGGGCAACGAGATGTGATTAACCAAGTCGGAAGCGCAATGTTTCTCCATTTAAAGT 936
QY      1019 TAAATTTTGTGTGATCCCAAGGCTCATTAACAATATTTTGTGATGTATTTGGC 1078
Db      937 GAGATGTTCTGGCTGATTAATAAAAGCCAAATACATTTATGATGACATTATAGCTGC 996
QY      1079 GGATCCCAATCACTGGAGGTTTCAAGGAGGCAAAATGGGTTCTTGGCGCAAGCGACAC 1138
Db      997 TGATGACTGTCTGTATTAATTTTCACAATTTCTGGTGTGATGTGGCTGTGAAGCCGACCC 1056
QY      1139 CAATGTGCAAGAAATCGGGTCTATATGATCCGGAATTCGCCCAACACTGGGGCTCACTG 1198
Db      1057 CGAAATGCG---CAAAAGAGATGTACATTCACCCGGACAGCCCGCTACTGGGGAACAGTG 1113

```

```

QY      1199 GATGGCGCAAGAAATCTCTTTGGAAATTTAAACTTACGAAACAACAAGAGCTTCAAA 1258
Db      1114 GATGTCCAAAGTGTGCTCACTTTCCACAAACTGAACCTACCAACCAACATTTTCAG----- 1166
QY      1259 TAACAATGGCGCAGATGTGTGTTTACAGTCTTGGACAAGTACCAAGCCCGCTGCATGT 1318
Db      1167 --ACAAACATGATTTATCTATTAATTTGAATCTCATGCAAAATACAGCCCGGTTCCACAT 1224
QY      1319 GGTGAAGTGAACGAGAGACGGCAGGAGACACTAGCCAGCCCGCGCGTGCAGAGCTT 1378
Db      1225 TGTAGAGCGCAATGACATCTTGAACCT-----CCCTTATAGTACATTTGCGACATA 1275
QY      1379 CACTTCCCTGAGACTGATTCAGTCCGCTGACCGCTACCGAAGACAGATATTACACA 1438
Db      1276 CTGTTCCTCCGAACTGATTTATGCTGTGTGACTGATACAGAAATGATATAGATAACCA 1335
QY      1439 ACTGAAAATAGATCAAAACCTTTTGCAAAAGATTTGCGGATTAAT 1484
Db      1336 GTTAAATATGACAAACACCTTTTGCAAAAGTTTCGGGACACT 1381

```

```

Search completed: December 20, 2005, 16:05:30
Job time : 548 secs

```


Db 61 CAGGACCGCTTTCATCGGCTGCACAAGACGAAAGATCAAAAAGTAGCCCTTTCTGATTCG 120
Qy 121 TGCATAGTGTCAATTGGCCCAATCTCTTCTCCAGGAGAAAAAAGTAATCAAACTT 180
Db 121 TGCATAGTGTCAATTGGCCCAATCTCTTCTCCAGGAGAAAAAAGTAATCAAACTT 180
Qy 181 TGAGAGAGATTTGCTGCTTGAAGTGTCTTCTGATGAGGGGGTCTGTGGATTTCTAG 240
Db 181 TGAGAGAGATTTGCTGCTTGAAGTGTCTTCTGATGAGGGGGTCTGTGGATTTCTAG 240
Qy 241 TTTATGATTAATAGACTTTAAAAACAAGGACGGGAGGCGAGTTCAGGTTCTAGAG 300
Db 241 TTTATGATTAATAGACTTTAAAAACAAGGACGGGAGGCGAGTTCAGGTTCTAGAG 300
Qy 301 CTATCAGCTGGAGCACTGCTCTTCTCTTCTATCATGCTCTCCAAAGAAATTTCTCAATG 360
Db 301 CTATCAGCTGGAGCACTGCTCTTCTCTTCTATCATGCTCTCCAAAGAAATTTCTCAATG 360
Qy 361 TGAGAGAGAGCTAACCAATTCAGGCGGATCGAGGCTTGTCTTGCACATCCCATTA 420
Db 361 TGAGAGAGAGCTAACCAATTCAGGCGGATCGAGGCTTGTCTTGCACATCCCATTA 420
Qy 421 TCTCGACCACTGACAACTTGAGAGAGTTCACTTTGAAAAAAATTAACAAGGGAGTGA 480
Db 421 TCTCGACCACTGACAACTTGAGAGAGTTCACTTTGAAAAAAATTAACAAGGGAGTGA 480
Qy 481 CGAATCAGTCAGATACAGCAATTTTCTGACTCCAAGGACTCACAGGGGAGCTTCAGA 540
Db 481 CGAATCAGTCAGATACAGCAATTTTCTGACTCCAAGGACTCACAGGGGAGCTTCAGA 540
Qy 541 GAAAGTAACTCTCTCTGTCTTGGAGCGGGGTCTGAGGCTTGTGACAGTTTCAATGAGT 600
Db 541 GAAAGTAACTCTCTCTGTCTTGGAGCGGGGTCTGAGGCTTGTGACAGTTTCAATGAGT 600
Qy 601 CTGCTGACAGATGCTACTCTCTCTCTCACTTCAGTCCAGCCAGCACAAGTCTGGGCACTGCTC 660
Db 601 CTGCTGACAGATGCTACTCTCTCTCTCTCACTTCAGTCCAGCCAGCACAAGTCTGGGCACTGCTC 660
Qy 661 CCAAGTCCATGTTTCCCGTACCCCGGCGAGACGAGCCGGCGACCCCGCTTCTCAATG 720
Db 661 CCAAGTCCATGTTTCCCGTACCCCGGCGAGACGAGCCGGCGACCCCGCTTCTCAATG 720
Qy 721 GGAGCCCTAGCCGCTCATATGCGCCACACCGGTCATCAACAAGGAGCTTCAACAGCC 780
Db 721 GGAGCCCTAGCCGCTCATATGCGCCACACCGGTCATCAACAAGGAGCTTCAACAGCC 780
Qy 781 TCTCTGCAACTCTTCGCGGAGGGATACCCACCGGCGGCTAACCCCTTACCAAGCAGT 840
Db 781 TCTCTGCAACTCTTCGCGGAGGGATACCCACCGGCGGCTAACCCCTTACCAAGCAGT 840
Qy 841 ACGGCCACTCTCAACCAAGAGCTCCGTTTCAACAGTTCTCTTCACCCAGCCGGGCTGG 900
Db 841 ACGGCCACTCTCAACCAAGAGCTCCGTTTCAACAGTTCTCTTCACCCAGCCGGGCTGG 900
Qy 901 TGCCTGGGCAAGACAGGTTGTAACGTGACAGGAGCCCTTTGGCTGAATTTCAACCGG 960
Db 901 TGCCTGGGCAAGACAGGTTGTAACGTGACAGGAGCCCTTTGGCTGAATTTCAACCGG 960
Qy 961 ACCAAACGAGATGATCATACCAACAAGAGGAGGAGCATGTTCTCTTTTAAAGTTTA 1020
Db 961 ACCAAACGAGATGATCATACCAACAAGAGGAGGAGCATGTTCTCTTTTAAAGTTTA 1020
Qy 1021 ACAATTTCTGCTCGATCCACGGCTCATTAACAATTTTTTGTGATGATTTTGGCGG 1080
Db 1021 ACAATTTCTGCTCGATCCACGGCTCATTAACAATTTTTTGTGATGATTTTGGCGG 1080
Qy 1081 ATCCCAATACCTGAGGTTTCAAGAGGAGGAATGGTTCTCTGGGCAAAAGGGACACA 1140
Db 1081 ATCCCAATACCTGAGGTTTCAAGAGGAGGAATGGTTCTCTGGGCAAAAGGGACACA 1140
Qy 1141 ATGTGCAAGGAATCGGCTCTATATGATCCGGAATCCCGCAACACTGGGGCTCACTGGA 1200
Db 1141 ATGTGCAAGGAATCGGCTCTATATGATCCGGAATCCCGCAACACTGGGGCTCACTGGA 1200

Qy 1201 TGGCCCAAGAAATCTCTTTGAAAAATTAAAACTTACGAACAAGAGGCTTCAATA 1260
Db 1201 TGGCCCAAGAAATCTCTTTGAAAAATTAAAACTTACGAACAAGAGGCTTCAATA 1260
Qy 1261 ACAATGGGACAGATGATGTTTACAGTCTTGCACAGTACCAAGCCCGCTGCATGG 1320
Db 1261 ACAATGGGACAGATGATGTTTACAGTCTTGCACAGTACCAAGCCCGCTGCATGG 1320
Qy 1321 TGGAAGTGAACGAGAGCGGACGAGAGACATAGCCAGCCCGGCGGTGCAGACTTCA 1380
Db 1321 TGGAAGTGAACGAGAGCGGACGAGAGACATAGCCAGCCCGGCGGTGCAGACTTCA 1380
Qy 1381 CTTCCTGGAGACTCAGTTACATGCGGTCACCGGCTACGAACAGAGATTTACACAC 1440
Db 1381 CTTCCTGGAGACTCAGTTACATGCGGTCACCGGCTACGAACAGAGATTTACACAC 1440
Qy 1441 TGAATAATGATCACAACTTTTGCAAAAAGATTTGAGATTAATTATGACAGATCTACA 1500
Db 1441 TGAATAATGATCACAACTTTTGCAAAAAGATTTGAGATTAATTATGACAGATCTACA 1500
Qy 1501 CCGGCTGTGATGATGAAACCGCTGACCCCTTGCACCAAGACTCGCGGCTCGCATTCG 1560
Db 1501 CCGGCTGTGATGATGAAACCGCTGACCCCTTGCACCAAGACTCGCGGCTCGCATTCG 1560
Qy 1561 TGGCCGGGGCCCGCTACGCAAGGCGGCTCTTCTGAGAGGACAGTGGTGAAGTCACT 1620
Db 1561 TGGCCGGGGCCCGCTACGCAAGGCGGCTCTTCTGAGAGGACAGTGGTGAAGTCACT 1620
Qy 1621 ACGCAAGGCGCGCTTCCACCCGAGGCGGAGCGGAGCCCGGAGCGAGTACGAGCCGA 1680
Db 1621 ACGCAAGGCGCGCTTCCACCCGAGGCGGAGCGGAGCCCGGAGCGAGTACGAGCCGA 1680
Qy 1681 GCGTGCAGCACACCAAGGAGCTGTGTGCGCGGACAGGCGGAGGACCCGGGCGGCTT 1740
Db 1681 GCGTGCAGCACACCAAGGAGCTGTGTGCGCGGACAGGCGGAGGACCCGGGCGGCTT 1740
Qy 1741 CGCGGCAACGCTGATTTGTGACGCGCGGCAACAACCGGCTTGAATTGCGGCGCTCGG 1800
Db 1741 CGCGGCAACGCTGATTTGTGACGCGCGGCAACAACCGGCTTGAATTGCGGCGCTCGG 1800
Qy 1801 ATGACACGCGCCACGAGACTTTCGCGGCAACGCGGCGACGCTGCTTACCGCGGCGG 1860
Db 1801 ATGACACGCGCCACGAGACTTTCGCGGCAACGCGGCGACGCTGCTTACCGCGGCGG 1860
Qy 1861 GCGTGAAGGCGCTGCGCTGAGGCTGAGGCTGACATGCGCGGCGCTCGGCTTACAG 1920
Db 1861 GCGTGAAGGCGCTGCGCTGAGGCTGAGGCTGACATGCGCGGCGCTCGGCTTACAG 1920
Qy 1921 CCGACCGGTGGGGCGCGGCGGCAAGCGGCGGCGGAGTACGAGGACCAAGTCGAGGCT 1980
Db 1921 CCGACCGGTGGGGCGCGGCGGCAAGCGGCGGCGGAGTACGAGGACCAAGTCGAGGCT 1980
Qy 1981 CGGTGCTGCTCTGTGCGCAACAGCGCGCGGCGGCGGCGGAGTACGAGGACCAATC 2040
Db 1981 CGGTGCTGCTCTGTGCGCAACAGCGCGCGGCGGCGGCGGAGTACGAGGACCAATC 2040
Qy 2041 CTTACTTGGGCGAGAGGCGGAGGCTGTGCGCGGAGGCTTGGCGTGCCTGCGCGGCG 2100
Db 2041 CTTACTTGGGCGAGAGGCGGAGGCTGTGCGCGGAGGCTTGGCGTGCCTGCGCGGCG 2100
Qy 2101 CCGCGAGGAGCGCAAGCGCAAGGACCTGTCCGATTCAGCTGAGTGAAGAGCGGCTCT 2160
Db 2101 CCGCGAGGAGCGCAAGCGCAAGGACCTGTCCGATTCAGCTGAGTGAAGAGCGGCTCT 2160
Qy 2161 CGATCAAGTCCATGACTCCAGCACTCGGAGATTTACAGACAGCCAAAGCGAGCGGA 2220
Db 2161 CGATCAAGTCCATGACTCCAGCACTCGGAGATTTACAGACAGCCAAAGCGAGCGGA 2220
Qy 2221 TCTGCGGCGCGACACGCGGCTGTCCGAGAGTTGTCCCGCTCAAGAGCGAGGTGCTGG 2280
Db 2221 TCTGCGGCGCGACACGCGGCTGTCCGAGAGTTGTCCCGCTCAAGAGCGAGGTGCTGG 2280

QY 2281 CCCAGCGGAGCTGCGAGAGAACTGCGCCAGAGCACTTAGCGGCTACTAGTGGCTCTCTCT 2340
 DB 2281 CCCAGCGGAGCTGCGAGAGAACTGCGCCAGAGCACTTAGCGGCTACTAGTGGCTCTCTCT 2340
 QY 2341 GCGACAGCTAGGCG 2400
 DB 2341 GCGACAGCTAGGCG 2400
 QY 2401 CCTCAGCTCTTCCCGAGCTCGCGCTCGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
 DB 2401 CCTCAGCTCTTCCCGAGCTCGCGCTCGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
 QY 2461 ATTGACCTCGATGCGCGCTCTGCGAGCGAATAGTGCAGCTCTCGAGGCTGATTTTAC 2520
 DB 2461 ATTGACCTCGATGCGCGCTCTGCGAGCGAATAGTGCAGCTCTCGAGGCTGATTTTAC 2520
 QY 2521 CTTTTTGAAGCGAGCTCTGCAATTAAGTCTCAACCGACCTTCAACTTCTCTGTAACCTT 2580
 DB 2521 CTTTTTGAAGCGAGCTCTGCAATTAAGTCTCAACCGACCTTCAACTTCTCTGTAACCTT 2580
 QY 2581 TTGGTTTCTCTACT 2640
 DB 2581 TTGGTTTCTCTACT 2640
 QY 2641 GTCTTCTCTTACT 2700
 DB 2641 GTCTTCTCTTACT 2700
 QY 2701 TGGGAGCTCTCTGCGAGCGAGCGAGCTCGAGCCGCGCAAGTCTGCGGCTCCGCACTTAC 2760
 DB 2701 TGGGAGCTCTCTGCGAGCGAGCGAGCTCGAGCCGCGCAAGTCTGCGGCTCCGCACTTAC 2760
 QY 2761 CATAGATGTTGACTCTAGAACCTGAGCCACCAGCGGCTCTTCTCTTATCCCGAGTG 2820
 DB 2761 CATAGATGTTGACTCTAGAACCTGAGCCACCAGCGGCTCTTCTCTTATCCCGAGTG 2820
 QY 2821 GAT 2880
 DB 2821 GAT 2880
 QY 2881 ATGAT 2940
 DB 2881 ATGAT 2940
 QY 2941 ACAAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
 DB 2941 ACAAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
 QY 3001 TCTCATACT 3060
 DB 3001 TCTCATACT 3060
 QY 3061 TAGAACAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
 DB 3061 TAGAACAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
 QY 3121 AATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
 DB 3121 AATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
 QY 3181 TAAAGG 3186
 DB 3181 TAAAGG 3186
 RESULT 2
 US-10-029-386-20697
 ; Sequence 20697, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 ; FILE REFERENCE: ABOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 20697
 ; LENGTH: 864
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC009487.3
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
 ; OTHER INFORMATION: NT HIT: 91142971, EVALUATE 0.00e+00
 ; OTHER INFORMATION: SWISSPROT HIT: Q16550, EVALUATE 0.00e+00
 ; OTHER INFORMATION: EST_HUMAN HIT: AI201654.1, EVALUATE 0.00e+00
 ; US-10-029-386-20697
 Query Match 27.1%; Score 864; DB 6; Length 864;
 Best Local Similarity 100.0%; Pred. No. 1,4e-230; Indels 0; Gaps 0;
 Matches 864; Conservative 0; Mismatches 0;

QY 1493 GATCTACACCGGCTGTGACATGACCGGCTGACCCCTCGCCCAAGACTGCGCGCTC 1552
 DB 1 GATCTACACCGGCTGTGACATGACCGGCTGACCCCTCGCCCAAGACTGCGCGCTC 60
 QY 1553 GCAGATCGTCCCGGCG 1612
 DB 61 GCAGATCGTCCCGGCG 120
 QY 1613 GAGCAACTACGCGAAGGCGCGCTTCACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1672
 DB 121 GAGCAACTACGCGAAGGCGCGCTTCACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 1673 GAGCGCGAGCGTGCAGACACCAACGAGGCTGTGCGCGCGAGCAGGCGCGAGACCGCGG 1732
 DB 181 GAGCGCGAGCGTGCAGACACCAACGAGGCTGTGCGCGCGAGCAGGCGCGAGACCGCGG 240
 QY 1733 GCGGCGCTGCGCGCAACGCTGTTGTGACCGCGCGCAACACCGGCTGAGACTTGGCGCG 1792
 DB 241 GCGGCGCTGCGCGCAACGCTGTTGTGACCGCGCGCAACACCGGCTGAGACTTGGCGCG 300
 QY 1793 CTCGCGCTATGACAGGCGCAACGAGCTTGGCGGCAACGCGCGAGCTGCTCTTACGC 1852
 DB 301 CTCGCGCTATGACAGGCGCAACGAGCTTGGCGGCAACGCGCGAGCTGCTCTTACGC 360
 QY 1853 GCGCGCGGCGTGAAGCGCTGCGCGAGCTGCGAGCTGCAAGCTGAGTGGCGCGCGCTCGG 1912
 DB 361 GCGCGCGGCGTGAAGCGCTGCGCGAGCTGCGAGCTGCAAGCTGAGTGGCGCGCGCTCGG 420
 QY 1913 CTACTACGCGACCGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1972
 DB 421 CTACTACGCGACCGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 QY 1973 GTGCGGCTGCGTGCCTGCTGCGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2032
 DB 481 GTGCGGCTGCGTGCCTGCTGCGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
 QY 2033 CGCAATCCCTACCTGCGGCGAGGAGCGAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCG 2092
 DB 541 CGCAATCCCTACCTGCGGCGAGGAGCGAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCG 600
 QY 2093 GCCCGGCGCGCGAGAGCGCGCAAGCGCAAGGAGCTGCTGCGATGCGAGC 2152
 DB 601 GCCCGGCGCGCGAGAGCGCGCAAGCGCGCAAGGAGCTGCTGCGATGCGAGC 660
 QY 2153 GCCCTCTCGATGACAGTGCATGACGCTCGAGCGAGCTGCGGAGATTTCAGCAGGCGCAAGCG 2212

```
Db      661  |||||GCCCCCTTCGATCAAGTCCATGACTCCAGCAGCACTCGGGGATTATACGAGCGCAAGCCAAACG 720
Qy      2213  GAGCGGATCTTCGCGCGCGGACACGCGCGTGTCCGAGATTGTCTCCCGCTCAAGAGCGA 2272
Db      721  GAGCGGATCTTCGCGCGCGGACACGCGCGTGTCCGAGATTGTCTCCCGCTCAAGAGCGA 780
Qy      2273  GGTGTGCGCCGAGCGGAGCTGGAGAAAGACTGCGCCCAAGACATTAAGGGTACTATGG 2332
Db      781  GTGTGTGCGCCGAGCGGAGCTGGAGAAAGACTGCGCCCAAGACATTAAGGGTACTATGG 840
Qy      2333  CTTTACTCTGCACAGCTAGGCCG 2356
Db      841  CTTTACTCTGCACAGCTAGGCCG 864

RESULT 3
US-10-363-345A-9001/c
; Sequence 9001, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9001
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 9001
US-10-363-345A-9001

Query Match      19.0%; Score 604; DB 8; Length 1026;
Best Local Similarity 80.5%; Pred. No. 9.9e-158;
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;

Qy      1490  CACGATCTACACCGGCTGTGATGAGACCGCTGACCGCCCTCGGCCCAAGACTGCGCGG 1549
Db      949  CAAATCTACACCGGCTGTGATGAGACCGCTGACCGCCCTCGGCCCAAGACTGCGCGG 890
Qy      1550  CTCGAGATCTGTGCGCGGCGCGCTACCGCATGCGCGGCTTTCTTCGACAGCAGTT 1609
Db      889  CTCGGAATCTGTACCGGAAACCGGCTACGCCATTAACCGACTCTTTCTTCAAAACCAATT 830
Qy      1610  CGTAGCAACTACGCGCAAGCGCGCTTCCACCCGGGCGCGGCGCGCGCGCGG 1669
Db      829  CGTAAACAATCTACGCGCAAAACCGGCTTCCACCGGAAACGC--GACGCGAAACCGAAACCGAA 772
Qy      1670  TACGGAACCGGACGCTGCGGACACCAACGCGGTGCTGTGCGCGGACGCGCGGAGAACCC 1729
Db      771  TACGGAACCGGACGCTGCGGACACCAACGACTACTATGCGCGGACCAAAACCGAAACCC 712
Qy      1730  GGGCGCGCGCTGCGCGCAACGCTGTTGTGACGCGCGGCAACACCGGCTGAGCTTGC 1789
Db      711  GAAACCGC--CTGCGCGCAACGCTATTATTATAGCGCGACCAACCAACGACTAACTTGC 653
Qy      1790  GGCCTGCGCTATGACACGCGCAACGAGCTTGGCGGCAACGCGCGCAAGCTGCTCTTA 1849
Db      652  GACTCTGACCTATTAACACGACCAACGAACTTGGCGAAACGCGACACGCTACTCTCTTA 593
Qy      1850  CGCGCGCGGCGGCGTGAAGCGCTGCGGTGAGGCTGAGGCTGAGCTGCGCGCGCT 1909
Db      592  CGCGAGCAGAACTGAAGAAACGCT--ACGCTACAAACTAACAACCTAACCGCGCT 534
Qy      1910  CGGCTACTACGCGGACCGCTGCGGCGCGCGCGGCGCGGAGTCCCGCGAGTACTGCGGAC 1969
|||
```

```
Db      533  CGACTACTACGCGGACCGCTGTGAATTAAGCCCGCAATCCCCCGCAATACTACGACAC 474
Qy      1970  CAAATCGGCTCGGTGTGCTGCTGTGCGCCCAACGCGCGCGCGCGCGCGCATGGC 2029
Db      473  CAAATCGAATCTGATCT--CCCTACTAACCAACGCGCGGACGCGCGCGCATTAAC 415
Qy      2030  CGGCGCAATCTCTTACTGTGGGCGAGAGGCGGCGCTGCGCGCGCGCGCGCGCT 2089
Db      414  CGAGCGCAATCTCTTACTTAACGAAACGAAACGAAACCTTAACCGCGGACGCTGCGCT 355
Qy      2090  GCGCGCGCGCGCGCGCGGACGCGCAACCGCAAGACCTGTCCGATTCCAGCTGATCGA 2149
Db      354  ACCGCCC--GCGCGCGCGCAAAACCGCAAAACCTTATCCGATTCCAACTAAATCGA 296
Qy      2150  GAGCGCGCTCTGATTCAGATTCATTCGACTCGAGGACTCGGGGATTATGAGCAGGCCAA 2209
Db      295  AAGCGCTCTCTGATCAAAATCATGACTCGCAAGACTCGAAATTTACGAACAAACCA 236
Qy      2210  GCGGAGCGGATCTGCGCGGCGGACACGCGCGTGTCCGAGATTGTCTCCCGCTCAAGAG 2269
Db      235  ACGAAACGAAATCTGCGCGGACCGACACGCGGATCCGAAATTCGTCGCGCTCAAAA 176
Qy      2270  CGAGTCTGCGCGCGGACGCGGACTGCGAAGAAAGAACTGCGCCCAAGGACTTACGCGCTA 2329
Db      175  CGAAATCTAAACCGCAAGAACTACGAAAAAACTACGCCAAAAACATTAACGACTACTA 116
Qy      2330  TGGCTTACTCTGCGACGCTAAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGGAGC 2389
Db      115  TAACTTACTCTGCGACGCTAAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGGAGC 56
Qy      2390  CCCAGCAGCGCTCTACAGGCTCTTCCCAAGCTCCGCTCCCAACTCTCTCTT 2443
Db      55  CCGAACCAACCCCTCAACAATCTTCCCAACTCGGCTCCCGCACTCTCTCTT 2

RESULT 4
US-10-363-345A-9002
; Sequence 9002, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9002
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 9002
US-10-363-345A-9002

Query Match      19.0%; Score 604; DB 8; Length 1026;
Best Local Similarity 80.5%; Pred. No. 9.9e-158;
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;

Qy      1490  CACGATCTACACCGGCTGTGATGAGACCGCTGACCGCCCTCGGCCCAAGACTGCGCGG 1549
Db      78  CAAATCTACACCGGCTGTGATGAGACCGCTGACCGCCCTCGGCCCAAGACTGCGCGG 137
Qy      1550  CTCGAGATCTGTGCGCGGCGCGCTACCGCATGCGCGCTTTCTTCGACAGCAGTT 1609
Db      138  CTCGGAATCTGTACCGGAAACCGGCTAGCCATTAACGAGCTTTCTTCAAAACCAATT 197
Qy      1610  CGTAGCAACTACGCGCAAGCGCGCTTCCACCCGGGCGCGGCGCGGCGCGCGCGG 1669
Db      198  CGTAAACAATCTACGCGCAAAACCGGCTTCCACCGGAAACGC--GACGCGAAACCGAAACCGAA 255
|||
```



```

QY 1670 TACGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGACGAGCGGAGAGCC 1729
DB 256 TACGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGACGAGCGGAGAGCC 315
QY 1730 GGGCGCGCGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1789
DB 316 GAGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 374
QY 1790 GCGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1849
DB 375 GAGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 434
QY 1850 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1909
DB 435 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 493
QY 1910 GCGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1969
DB 494 GCGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 553
QY 1970 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2029
DB 554 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 612
QY 2030 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2089
DB 613 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 672
QY 2090 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2149
DB 673 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 731
QY 2150 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2209
DB 732 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 791
QY 2210 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2269
DB 792 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 851
QY 2270 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2329
DB 852 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 911
QY 2330 TGGCTTCTGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2389
DB 912 TGGCTTCTGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 971
QY 2390 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2443
DB 972 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1025

RESULT 5
US-10-363-483A-9001/c
; Sequence 9001, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9001
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
OTHER INFORMATION: Cpg-island No: 9001
US-10-363-483A-9001
Query Match 19.0%; Score 604; DB 9; Length 1026;
Best Local Similarity 80.5%; Pred. No. 9.9e-158;
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;

QY 1490 CAGCATCTACGCGGCTGTGACATGAGCGGCTGACCGGCGGAGCGGAGAGCC 1549
DB 949 CAAATCTACGCGGCTGTGACATGAGCGGCTGACCGGCGGAGAGCC 890
QY 1550 CTGCGAGATCTGCGCGGCGGCGGCTGACCGGCTGACCGGCGGAGAGCC 1609
DB 889 CTGCGAGATCTGCGCGGCGGCGGCTGACCGGCTGACCGGCGGAGAGCC 830
QY 1610 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1669
DB 829 GGTAAAGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 772
QY 1670 TACGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1729
DB 771 TACGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 712
QY 1730 GGGCGCGCGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1789
DB 711 GAGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 653
QY 1790 GCGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1849
DB 652 GAGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 593
QY 1850 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1909
DB 592 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 534
QY 1910 GCGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 1969
DB 533 GAGCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 474
QY 1970 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2029
DB 473 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 415
QY 2030 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2089
DB 414 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 355
QY 2090 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2149
DB 354 GCGCGCGGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 296
QY 2150 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2209
DB 295 GAGTCTGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 236
QY 2210 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2269
DB 235 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 176
QY 2270 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2329
DB 175 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 116
QY 2330 TGGCTTCTGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2389
DB 115 TGGCTTCTGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 56
QY 2390 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2443
DB 55 GCGGAGCGGAGCGGTGCGGACACCAAGGCGTCTGTGCGCGGAGAGCC 2

```

RESULT 6
US-10-363-483A-9002
; Sequence 9002, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9002
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 9002
US-10-363-483A-9002

Query Match 19.0%; Score 604; DB 9; Length 1026;
Best Local Similarity 80.5%; Pred. No. 9,9e-158;
Matches 768; Conservative 0; Mismatches 180; Indels 6; Gaps 5;

QY 1490 CACGATCTACACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCGCCGG 1549
DB 78 CAAATCTACACCGGTGTGACATGAGACCGCTGACCCCTGCGCCCAAGACTCGCCGG 137
QY 1550 CTCGAGATCTGACCGCGGGGCGCGCTACCGCATGCGCGCGCTCTTCTCTGACGACGATT 1609
DB 138 CTCGCAATCTGACCGCGGAAACCGCTACCGCATGCGCGCTCTTCTCTGACGACGATT 197
QY 1610 CGTGAACAATACCGCAAGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCG 1669
DB 198 CGTAAACAATACCGCAAGCGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCG 255
QY 1670 TACGACCGGCG 1729
DB 256 TACGACCGGCG 315
QY 1730 GAGCG 1789
DB 316 GACCG 374
QY 1790 GGCCTGCGCTATGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1849
DB 375 GACCTGCGCTATGACACG 434
QY 1850 CG 1909
DB 435 CG 493
QY 1910 CGGCTACTACG 1969
DB 494 CGGCTACTACG 553
QY 1970 CAACTGCG 2029
DB 554 CAACTGCG 612
QY 2030 CGGCG 2089
DB 613 CGGCG 672
QY 2090 GCG 2149
DB 673 GCG 731
QY 2150 GACG 2209

DB 732 AACGCCCTCTCGATCAATTCATCGACTCCACGACTCGAAAAATTACGAACAAACCA 791
QY 2210 GCGGAGCGGATCTCG 2269
DB 792 AGAAAAAGAACTCTCG 851
QY 2270 CGAGGTCTGCG 2329
DB 852 CGAAATCTAACCGAAAGAACTCGAAAAAACTAGCGCGCGCGCGCGCGCGCGCGCG 911
QY 2330 TGGCTTCTACTCGACAGCTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2389
DB 912 TAACTTCTACTCGACAGCTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 971
QY 2390 CCCAGCG 2443
DB 972 CCCAACCAACCGCTCACAATCTTCCCAACTCCGCGCTCCCAACACTCTCTCTT 1025

RESULT 7
US-10-029-386-6986
; Sequence 6986, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Penn, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOmica-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6986
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009487.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EST HUMAN HIT: H10108.1, EVALU0 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q16650, EVALU0 5.00e-15
; OTHER INFORMATION: NT HIT: g111429971, EVALU0 0.00e+00
US-10-029-386-6986

Query Match 16.2%; Score 515; DB 6; Length 515;
Best Local Similarity 100.0%; Pred. No. 5,6e-133; Indels 0; Gaps 0;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2236 GCGCGGTCTCGAGAGTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2295
DB 1 GCGCGGTCTCGAGAGTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 2296 AGAAGAACTGCG 2355
DB 61 AGAAGAACTGCG 120
QY 2356 CCCCTGCG 2415
DB 121 CCCCTGCG 180
QY 2416 CGAGTCTGCG 2475
DB 181 CGAGTCTGCG 240
QY 2476 GCGGTCTGCG 2535

Db 241 GCCGTGACGGAATTAAGTGAAGTCTCCGAGCTGATTTAACTTTTGGACACA 300
Qy 2536 GTCTGTGAATTAAGTCAACCGACTTCTCACTTTGTGTAACCTTTTGTCTTACTT 2595
Db 301 GTCTGTGAATTAAGTCAACCGACTTCTCACTTTGTGTAACCTTTTGTCTTACTT 360
Qy 2596 ACTCTTCTTGTGAGATTATCTCTCAATTTCCCTCCCTCTGTCTTCTTACTT 2655
Db 361 ACTCTTCTTGTGAGATTATCTCTCAATTTCCCTCCCTCTGTCTTCTTACTT 420
Qy 2656 CCTACTTCTTCTTGTGTAATTAACCTTCACTTTAGAGACCTGGGAGTCTGTCA 2715
Db 421 CCTACTTCTTCTTGTGTAATTAACCTTCACTTTAGAGACCTGGGAGTCTGTCA 480
Qy 2716 GGCAGCAGGATTCGACCCGCGCAAGTCTGGGCT 2750
Db 481 GGCAGCAGGATTCGACCCGCGCAAGTCTGGGCT 515

RESULT 8
US-10-842-072-37/c
; Sequence 37, Application US/10842072
; Publication No. US20050158731A1
; GENERAL INFORMATION:
; APPLICANT: PLASS, CHRISTOPH
; TITLE OF INVENTION: DETECTION OF METHYLATED CPG RICH SEQUENCES DIAGNOSTIC FOR
; FILE REFERENCE: 22727-04242
; CURRENT APPLICATION NUMBER: US/10/842,072
; CURRENT FILING DATE: 2004-05-10
; PRIOR APPLICATION NUMBER: 09/775,398
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 37
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-842-072-37

Query Match 16.1%; Score 514; DB 9; Length 514;
Best Local Similarity 100.0%; Pred. No. 1.1e-132; Indels 0; Gaps 0;

Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1504 GCTGTGACATGACCGCTGACCCCTCGCCCAACGACTCGCGCTGCAAGATGTC 1563
Db 514 GCTGTGACATGACCGCTGACCCCTCGCCCAACGACTCGCGCTGCAAGATGTC 455
Qy 1564 CCGGGGCGGCTTACCGCATGCGCGGCTCTTCTGTGAGACAGTTCGTGACACTACG 1623
Db 454 CCGGGGCGGCTTACCGCATGCGCGGCTCTTCTGTGAGACAGTTCGTGACACTACG 395
Qy 1624 CCAAGCGCGCTTACCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1683
Db 394 CCAAGCGCGCTTACCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
Qy 1684 TGCCTGACACCAACGAGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1743
Db 334 TGCCTGACACCAACGAGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275
Qy 1744 CCGCAACGCTGTTGTGACCGCGCGCAACACCGCTGACCTTTCGCGGCTTCG 1803
Db 274 CCGCAACGCTGTTGTGACCGCGCGCAACACCGCTGACCTTTCGCGGCTTCG 215
Qy 1804 ACACGCGCAACGAGCTTTCGCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 1863
Db 214 ACACGCGCAACGAGCTTTCGCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 155
Qy 1864 TGAAGCGCTGCGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCT 1923
Db 154 TGAAGCGCTGCGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCT 95
Qy 1924 ACCCGTCGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1983

Db 94 ACCCGTCGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 35
Qy 1984 TGTGCGCTGCTGGCGCCCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2017
Db 34 TGTGCGCTGCTGGCGCCCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1

RESULT 9
US-10-363-345A-9003
; Sequence 9003, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO: 9003
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-363-345A-9003

Query Match 14.7%; Score 469.6; DB 8; Length 1026;
Best Local Similarity 71.7%; Pred. No. 4.2e-120; Indels 6; Gaps 5;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;

Qy 1491 ACAGCTACACCGGCTGTGACATGACCGCGCTGACCCCTCGCCCAACGACTCGCGCGC 1550
Db 79 ACAGCTACACCGGCTGTGACATGACCGCGCTGACCCCTCGCCCAACGACTCGCGCGC 138
Qy 1551 TCGCAGATCGGCG 1610
Db 139 TCGCAGATCGGCG 198
Qy 1611 GTGACACATGACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1670
Db 199 GTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 256
Qy 1671 ACAGCGGAGCGTCCGCGCAACGAGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCG 1730
Db 257 ACAGCGGAGCGTCCGCGCAACGAGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCG 316
Qy 1731 GCGCGCGCTCGCGCAACGAGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1790
Db 317 GCGCGCGCTCGCGCAACGAGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 375
Qy 1791 GCGTGGCTTATGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1850
Db 376 GCGTGGCTTATGACACG 435
Qy 1851 GCGGCG 1910
Db 436 GCGGCG 494
Qy 1911 GCGTACTACG 1970
Db 495 GCGTACTACG 554
Qy 1971 AAGTGGGCTGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2030
Db 555 AAGTGGGCTGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 613
Qy 2031 GCGCGCAATCCCTACCTGGGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2090

Db 614 GGCCTTAATTTTATTTGGGCGAGGAGGTCGAGGGTTTGTCTGTCGAGGTTTCTGTCGTTG 673
Qy 2091 CCGCCCGGCGCCCGGAGGAGCCCAAGCCCAAGACCTGTCCGATTCCAGCTGATCCAG 2150
Db 674 TCGTTC-GCGTCGTGAGAGACGTTAAGTTTGAAGTTTGTTCGATTATTAAGTTGATCAG 732
Qy 2151 ACGCCCTCTGTCAGATCCATCGACTCCAGCTCGGGATTTTACAGAGGAGCAAG 2210
Db 733 ACGTTTTCATTCATTAAGTTTATCGATTTCAGATTTCGGGATTTACAGATAGTTAAG 792
Qy 2211 CGAGCGGAGATCTCGCCGCGGAGCAAGCCGCTGTCGAGAGTTCTCCCGCTCAAGAGC 2270
Db 793 CGAGGCGGATTTTCGTCGCTGATAGTTTCGTTTCAGAGATTCGTTTTCGTTTAAAGAGC 852
Qy 2271 GAGGTGCTGCGCCGAGGAGCTGCGAGAGAACTGCGGCAAGATTAAGCGCTACTAT 2330
Db 853 GAGGTGTTGGTTTAAAGGAGATTGCGAGAAATTTGGTTTAAAGAAATTAAGCGTTATAT 912
Qy 2331 GGCCTTCTACGACAGTAGAGCCGCGCCCTGCGCCGCGCCGCGCGCGAGAGCC 2390
Db 913 GGTTTTATTCGTAATTAAGTTTATAGTTTTCGTTTTCGTTTCGTTTCGTTTCGAGATT 972
Qy 2391 CGAGCGAGCCCTCAAGCTCTTCCGAGCTCGGCTCCCACTCTCTCTTG 2444
Db 973 TTAGTTAGTTTATTAAGTTTATTAAGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCG 1026

RESULT 10
US-10-363-345A-9004/C
; Sequence 9004, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9004
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 9004
US-10-363-345A-9004

Query Match 14.7%; Score 469.6; DB 8; Length 1026;
Best Local Similarity 71.7%; Pred. No. 4.2e-120;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;
Qy 1491 ACGATCTACACCGCTGTGATGACATGACCGCTGACCCCTCGCCCAAGACTCGCGAGC 1550
Db 948 AAGATTATATGCTGTGTATGATGATCGTTGATTTTTCGTTTAAACATTCGTCGCT 889
Qy 1551 TCGCAGATCTGCGCCGCGGCGCTAAGCCATGCGCGCTCTTCTCGAGAGCAAGTTC 1610
Db 888 TCGTAGATCGTGTTCGGGGTTCGTTACGTTATGTCGTTTTCGTTTTCGAGATTAGTTC 829
Qy 1611 GTGAGCACTACCGCAAGCCGCTTCACACCGGCGCGGCGCGGCGCGCGCGCT 1670
Db 828 GTGAGTAATTAAGTTTAAAGTTTCGTTTTCGTTTTCGAGGCGCG--CGCGGTTTCGAGTTCG 771
Qy 1671 ACGGACCGAGGCTGCGGAGCAAGAGGCTGTGTCGCGAGAGGCGAGAGAGCCG 1730
Db 770 ACGGATCGTAGCGTGTGTAATTAAGGAGTTGTGTCTGTAAGATCGAGATTTCG 711
Qy 1731 GCGCGCGCTCGCGCGAGCGCTGTGTAAGCGCGCGCGCAACACCGCTGAGCTTCGCG 1790
Db 710 GCGCGGCTTCGTCGTAAGCGTTGTTGAGCGTGTGTAATTAAGTTCGTTTCGCG 652

Qy 1791 GCTTCGCTTATGACAGGCGCAGAGCTTCGCGGAGCAAGCGGCGCAGCTCTCTTAC 1850
Db 651 GTTTCGTTTATGATACGATGATTCGAGTTTCGCGGATTAAGCGGATTCGTTTTC 592
Qy 1851 GCGCGCGGCGGCTGAAGGCGCTGCGCTGAGCTGACAGCTGACATGCGCGCGCTC 1910
Db 591 GCGCGCGGCGGCTGAAGGCGCTGCGCTGAGCTGACAGCTGACATGCGCGCGCTC 533
Qy 1911 GCTTACTACGCGGACCGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCTGCGGAGC 1970
Db 532 GGTATTAACGTCGATTCGTCGAGTTGCGGCTGCTGATGTTTTCGATGATTCGATTC 473
Qy 1971 AAGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2030
Db 472 AAGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 414
Qy 2031 GCGCCCAATCCCTTACTGCGGAGAGAGCGCGAGGCTGCGCGCGAGCGCTGCGCTG 2090
Db 413 GCGCTTATTTTATTTTCGCGGAGAGAGGTCGAGGCTTTCGTCGAGCGCTTCGCTG 354
Qy 2091 CCGCCCGGCGCGCGCGAGAGAGCGCAAGCCCAAGACCTGTCCATTCAGCTGATGAG 2150
Db 353 TCGTTC-GCGTCGTGAGAGAGTTAAGTTTGAAGTTTTCGATTTTGAAGTTCGAG 295
Qy 2151 ACGCCCTCTGATCAAGTCCATGCACTCCAGCACTCGGAGATTTACAGAGAGCCAG 2210
Db 294 ACGTTTTCGATTAAGTTTATGATTTTACGATTTTCGAGATTTTCAGATGATTAAG 235
Qy 2211 CGAGCGGAGATCTCGCGCGGAGAGCGCCGCTGTCCAGAGTTGTCGCCCTCAAGAGC 2270
Db 234 CGAGGCGGATTTTCGTCGCTGATAGCTGTTCGATGAGATTCGTTTTCGTTTAAAGAGC 175
Qy 2271 GAGGTGCTGCGCCGAGGAGCTGCGAGAAATCTGCCCAAGACATTAAGCGCTACTAT 2330
Db 174 GAGGTGTTGTTTACCGGATTCGAGAAAGATTCGTTAAGATTAAGCGGTTATAT 115
Qy 2331 GGCCTTCTACGACAGTAGAGCCGCGCTGCGCGCGCGCGCGCGCGCGAGAGCC 2390
Db 114 GGTTTTATTCGTAATTAAGTTTATGATTTTTCGTTTTCGTTTCGTTTCGAGATT 55
Qy 2391 CGAGCGAGCCCTCAAGCTCTTCCCAAGCTCGGCTCCCACTCTCTCTTG 2444
Db 54 TTAGTTAGTTTATTAAGTTTATTAAGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCG 1

RESULT 11
US-10-363-483A-9003
; Sequence 9003, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE OF INVENTION: illnesses
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9003
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 9003
US-10-363-483A-9003

Query Match 14.7%; Score 469.6; DB 9; Length 1026;
Best Local Similarity 71.7%; Pred. No. 4.2e-120;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;

```

QY 1491 AGCATCTACACCGCTGTGACATGACACCGCTGACCCCTCGCCCAAGACTGCGCGGC 1550
DB 79 AGGATTTATATCGTTGATATGATGATGATGATTTTTCGTTTAACGATTCGTCGGGT 138
QY 1551 TCGCATCTGCGCCCGGGGCGCGCTACGCGCATGCGCGCTCTTCTCGAGGACAGTTTC 1610
DB 139 TCGTATGATCGTTGCGGGGTGTTACGTTATGTCGCTTTTCTTTTGTAGGATTAAGTTTC 198
QY 1611 GTGACCAACTACGCAAGCGCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGGT 1670
DB 199 GTGAGTATTAAGTTAAGTTGTTTATTTTTCGCGCGCGCGCGCGCGCGCGGT 256
QY 1671 ACGGACCGGAGCGTGTGCGGACACCAACGCGCGCTGTGTGCGCGGACGCGCGCGCGCG 1730
DB 257 ACGGATCTGAGCGTGTGCTATTAACGCGGTTGTTGTCTGATGAGGTTCAGGATTCG 316
QY 1731 GCGCGCGCGCGCGCGCGCGCGCTGTTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1790
DB 317 GCGCGCG-TTTTGTCGTACGTTGTTGTTGAGCGTGTGCGGTTAATATCGGTTTTCGCG 375
QY 1791 GCGTGGCTATGACACGCGGACACGCGACTTCGCGCGCGCGCGCGCGCGCGCTCTTAC 1850
DB 376 GTTTCGCTTATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 435
QY 1851 GCGCGCGCGCGCGCGCGCGCGCTGCGCTGCGGACGCTGCGGACGCTGCGCGCGCGCG 1910
DB 436 GCGCGCGCGCGCGCGCGCGCGCTG- CGTTGATGTTGATGTTGATGTTGATGTTGATG 494
QY 1911 GCGTACTACGCGGACCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1970
DB 495 GGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
QY 1971 AAGTGGGCTCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2030
DB 555 AAGTGGGCTCGGTCG- TTTTGTGTTGTTTAAAGTGTGCGCGCTGCTGCGGTAATGTC 613
QY 2031 GCGCGCAATTCCTACCTGCGCGGAGAGCGCGGAGCGCTGCGCGCGCGCGCGCGCGCG 2090
DB 614 GCGGTTAATTTTATTTTGGGCGAGGAGTGAAGGTTGTTGTTGTTGTTGTTGTTGTTG 673
QY 2091 CCGCGCGCGCGCGCGCGGACCGCGGACCGCGGACCGCGGACCGCGGACCGCGGACCG 2150
DB 674 TCGTTC- GCGTCTCTGAGGACGTTAAGTTTAAAGTTTGTTCGATTTTAAAGTTGAG 732
QY 2151 AGCCCTCTCTGATCAAGTCTGATCACTGCGGAGCTGCGGAGTTTACGAGCGCGCAAG 2210
DB 733 AGCTTTTTCGATTAAGTTTATGATTTTAAAGGATTTTACGAGTTTAAAGTTTAAAG 792
QY 2211 CCGAGGCGGATCTGCGCGCGCGGACGCGCGCGGTGCGGAGTTTCTGCGCGCTCAAG 2270
DB 793 CCGAGGCGGATTTGTCGCGCGGATCGTTGATGTTGTTGTTGTTGTTGTTGTTGTTG 852
QY 2271 GAGTGTGCGCGGACGCGGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2330
DB 853 GAGTGTGTTGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 912
QY 2331 GCGTCTCTGATCAAGTCTGATCACTGCGGAGCTGCGGAGTTTCTGCGCGCTCAAG 2390
DB 913 GGTTTTATTCGATTAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 972
QY 2391 CCGAGGCGGCTCTGCGGAGCTCTGCGGAGCTCTGCGGAGCTCTGCGGAGCTCTGCG 2444
DB 973 TTAGTTAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 1026

```

RESULT 12
 US-10-363-483A-9004/c
 ; Sequence 9004, Apollication US/10363483A
 ; Publication No. US2005064401A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexander Olek
 ; APPLICANT: Christian Plepenbrock
 ; APPLICANT: Kurt Berlin

```

; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 9004
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 9004
US-10-363-483A-9004

Query Match      14.7%; Score 469.6; DB 9; Length 1026;
Best Local Similarity 71.7%; Pred. No. 4.2e-120;
Matches 684; Conservative 0; Mismatches 264; Indels 6; Gaps 5;

QY 1491 AGCATCTACACCGCTGTGACATGACACCGCTGACCCCTCGCCCAAGACTGCGCGGC 1550
DB 948 AGGATTTATATCGTTGATATGATGATGATGATTTTTCGTTTAACGATTCGTCGGGT 889
QY 1551 TCGCATCTGCGCCCGGGGCGCGCTACGCGCATGCGCGCTCTTCTCGAGGACAGTTTC 1610
DB 888 TCGTATGATCGTTGCGGGGTGTTACGTTATGTCGCTTTTCTTTTGTAGGATTAAGTTTC 829
QY 1611 GTGACCAACTACGCAAGCGCCGCTTCCACCCGCGCGCGCGCGCGCGCGCGCGGT 1670
DB 828 GTGATTAATTAAGTTAAGTTTATGTTTATTCGCGCGCG- CGCGGTTGCGGTCGGGT 771
QY 1671 ACGGACCGGAGCGTGTGCGGACACCAACGCGCGCTGTGTGCGCGGACGCGGACCGG 1730
DB 770 ACGGATCTGAGCGTGTGCTATTAACGCGGTTGTTGTGCTGATGAGTTCAGGATTCG 711
QY 1731 GCGCGCGCGCGCGCGGACCGCGGAGCTGTTGAGCGCGCGCAACGCGGCTGACTTTCG 1790
DB 710 GCGCGCG- TTTGTCGTAAACGTTGTTGTTGAGCGTGTGTTAATCGGTTGATTTTCG 652
QY 1791 GCGTGGCTATGACACGCGGACCGGACTTCGCGGAGGAGCGCGGACCGCGGCTCTTAC 1850
DB 651 GTTTCGTTTATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 592
QY 1851 GCGCGCGCGCGGAGGAGCGCTGCGCTGCGGACGCTGCGGACGCTGCGGACGCTGCG 1910
DB 591 GCGCGCGCGCGGAGGAGCGCTG- CGTTGATGTTGATGTTGATGTTGATGTTGATG 533
QY 1911 GCGTACTACGCGGACCGCGCTGCGGAGCGCGGACCGCGGACCGCGGACCGCGGACCG 1970
DB 532 GGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 473
QY 1971 AAGTGGGCTCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2030
DB 472 AAGTGGGTCGCTG- TTTTGTGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 414
QY 2031 GCGCGCAATTCCTACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2090
DB 413 GCGGTTAATTTTATTTTGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354
QY 2091 CCGCGCGCGCGCGCGGAGGAGCGCAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2150
DB 353 TCGTTC- GCGTCTCTGAGGACGTTAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 295
QY 2151 AGCCCTCTCTGATCAAGTCTGATCACTGCGGAGCTGCGGAGTTTACGAGCGCGCAAG 2210
DB 294 AGCTTTTTCGATTAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 235
QY 2211 CCGAGGCGGATCTGCGCGCGGACGCGCGGTGTCGAGAGTTTCTGCGCGCTCAAGAG 2270
DB 234 CCGAGGCGGATTTGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 175
QY 2271 GAGTGTGCGCGGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2330

```

Db 174 GAGGTGTTGTTAGTTCGAGAAAGATTGCTTAAAGATTAAGCGTTATAT 115
Qy 2331 GAGCTTACTGACAGCTAGAGCCCTGCTCCCGCCCGCCCGCCGAGACC 2390
Db 114 GGTTTTATGATAGTATAGTTCGTTTGTCTTGGTTCGTCGCGGTGCGATT 55
Qy 2391 CCAGCAGCCCTCAGCTCTCCCGAGCTCCGCTCCCACTCTCTCTG 2444
Db 54 TTAGTTAGTTTATAGTTTATTTTATTTTATTTTATTTTATTTTATTTT 1

RESULT 13

US-09-918-995-2491
; Sequence 2491, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2491
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-2491

Query Match 13.2%; Score 419; DB 3; Length 457;
Best Local Similarity 98.8%; Pred. No. 3.9e-106;

Matches 422; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2299 AGAATGCGCCAGGACATTAAGCGCTACTAGCTTACTCGACAGCTAGCGCC 2358
Db 31 AAACCTGGCCCAAGACATTAAGCGCTACTAGCTTACTCGACAGCTAGCGCC 90
Qy 2359 CTGCGCGCCGCGCCGCGCGCGCGCGAGCCCGCAGCCGCTTCTTCCCA 2418
Db 91 CTACCGCGCGCGCGCGCGCGCGCGAGCCCGCAGCCGCTTCTTCCCA 150
Qy 2419 GCTCGCGCTCCCGCACTCTCTCTGCGGACCACTATTATTTGACCTCGATGCGC 2478
Db 151 GCTCGCGCTCCCGCACTCTCTCTGCGGACCACTATTATTTGACCTCGATGCGC 210
Qy 2479 GCTCGAGGAAATAGTGAAGTCTCCGAGCGTGAATTTTAACTTTTGAAGAGTGC 2538
Db 211 GCTCGAGGAAATAGTGAAGTCTCCGAGCGTGAATTTTAACTTTTGAAGAGTGC 270
Qy 2539 TCTGCAATTAAGCTCAGCAGCTTCACTTTGCTGAACCTTTGTTTCTTACTTACT 2598
Db 271 TCTGCAATTAAGCTCAGCAGCTTCACTTTGCTGAACCTTTGTTTCTTACTTACT 330
Qy 2599 CTCTCTCTGAGATTAATCTCTCTCAATTTCCCTCTCTCTCTCTCTCTCTCTCT 2658
Db 331 CTCTCTCTGAGATTAATCTCTCTCAATTTCCCTCTCTCTCTCTCTCTCTCTCTCT 390
Qy 2659 ACTTCCTCTCTGATGAATGAATCTCTCACTTTGAGAGAGCTGGAGAGTCTCTGAGGC 2718
Db 391 ACTTCCTCTCTGATGAATGAATCTCTCACTTTGAGAGAGCTGGAGAGTCTCTGAGGC 450
Qy 2719 AGCAGCG 2725
Db 451 AGCAGCG 457

RESULT 14

US-10-723-860-2746
; Sequence 2746, Application US/10723860
; Publication No. US20040253606A1

; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnick, Albert
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUB01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2746
; LENGTH: 2061
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-2746

Query Match 10.9%; Score 346; DB 8; Length 2061;
Best Local Similarity 74.9%; Pred. No. 2.5e-85;
Matches 433; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 912 GCACAGGTGTAAGCTGTGCAACAGGCGCTTGTGCTGAATTTACCGGCACCAACGAG 971
Db 799 GCCACGCTAAGCTGTGCAACAGGCGCTGTGCTGAATTTACCGGCACCAACGAG 858
Qy 972 ATGATCATCAACCAACAGGAGGCGATGTTCTTTTAAAGTTTAACTTCTGAT 1031
Db 859 ATGATCATCAACCAACAGGAGGCGATGTTCTTTTAAAGTTTAACTTCTGAT 918
Qy 1032 CTGATCCCAAGGCTGATTAATTTTGTGATGATGATTTTGGCGATCCCAATCAC 1091
Db 919 CTCAACCCACCGCCACTACATGTTTGTGGAATGTTCTGCGCCACCTTAACAC 978
Qy 1092 TGAAGTTTCAAGGAGCAATGAGTCTTTGCGGCAACGCAACCAATGTCAGAA 1151
Db 979 TGGCGCTTCAAGGAGCAATGAGTCTTTGCGGCAACGCAACCAATGTCAGAGC 1038
Qy 1152 AATGGGTCTATATGATCTCGGATTTCCCGCAACCTGGGGCTCACTGAGAGCCCAAGAA 1211
Db 1039 AATGAAATGATATGATCTCAACGAGATCTTAATCTGATCTTCCCACTGATGAGCAGAG 1098
Qy 1212 ATCTCTTTGGAATTAATTAATTAAGCAACCAACCAAGAGCTTCAATTAATTAATGAGCAG 1271
Db 1089 ATTTCAATTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1158
Qy 1272 ATGATGTTTAAAGTCTTGTGCAACAGTACCAAGCCCGCTGATGTTGGAATGAAC 1331
Db 1159 ATGATGTTTAAAGTCTTGTGCAACAGTACCAAGCCCGCTGATGTTGGAATGAAC 1218
Qy 1332 GAGAGCGCAGGAGGACACTAGCAGCAGCCCGCGCGGTGAGAGTCTTCTCCCTGAG 1391
Db 1219 GAGAGTGGCGTGAAGACTTGAATGAGCCTCAAGAGCCCAAGCTTTTACTTCTTCAAGAA 1278
Qy 1392 ACTGATTCATGCGCGTCAACCGCTCAACAGAAACAGGATTTACAACTGAATAATAGAT 1451
Db 1279 ACGAATTCATGAGAGTACGCTTACCAAAACCGATATTAATCACTTAATTAATTAATTAATTAAT 1338
Qy 1452 CACAACCTTTTGCAAAAGATTTGCGGATTAATTAATGA 1489
Db 1339 CATTAACCTTTTGCAAAAGGCTTCAAGAGACAACTATGA 1376

RESULT 15

US-10-723-860-6839
; Sequence 6839, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnick, Albert
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &

This Page Blank (uspto)